



Thomas, P., & Vaughan, GJ. (2015). Pitfalls in the application of utility functions to the valuation of human life. *Process Safety and Environmental Protection*, 98, 148-169.
<https://doi.org/10.1016/j.psep.2015.07.002>

Publisher's PDF, also known as Version of record

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.psep.2015.07.002](https://doi.org/10.1016/j.psep.2015.07.002)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Process Safety and Environmental Protection

journal homepage: www.elsevier.com/locate/psep

Pitfalls in the application of utility functions to the valuation of human life

P.J. Thomas*, G.J. Vaughan

Safety Systems Research Centre, Queen's School of Engineering, Bristol University, Queen's Building, University Walk, Clifton BS8 1TR England

ARTICLE INFO

Article history:

Received 27 May 2015

Accepted 2 July 2015

Available online 14 July 2015

Keywords:

Safety

Utility functions

Value of a Prevented Fatality

VPF

Health and Safety Executive

Human life

ABSTRACT

Safety strategies in the process and other industries depend ultimately on how much the owners and operators decide should be spent on protection systems to protect workers and the public from potential plant hazards. An important input to decisions of this sort is the value of life, which needs to be assessed in a valid manner so that safety decisions can be made properly. A key reference point for decisions on safety investment decisions in the UK is a 1999 study on the “value of a prevented fatality” (VPF), which employs a two-injury chained model that has been shown previously by the present authors to possess internal inconsistencies. The 1999 study made extensive use of utility functions to interpret survey data, and it is this feature that is explored in this paper. It will be explained here how different forms of utility function of the Exponential family can produce the same figure for an intermediate parameter in the calculation of the VPF from the two-injury chained model. Exponential utility functions are, however, unlikely to provide a realistic representation if their calculated risk-aversions need to be negative or zero in order to match survey data, which would imply an incautious attitude amongst those taking decisions on safety. The use of an incompletely specified wealth threshold in the utility modelling is explored in the light of a proposal by the authors of the 1999 study that a second utility function can be used to determine the individual's utility when his wealth lies below the threshold, which constitutes the lower limit of validity of the first utility function. The proposition is shown to be untenable. The results presented in this paper raise further concerns about the lack of validity of the 1999 study on which the UK VPF is based and hence on the safety decisions that have been made in consequence.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Safety decisions in many industrial situations, particularly in high-hazard industries such as nuclear, oil and gas and chemicals, depend ultimately on a consideration of how much should be spent on safety measures to protect workers and the public from potential hazards. In the UK the legal requirement is to do all that is reasonably practicable to ensure the health and safety of workers and the public. This requirement means that it is necessary to compare the sacrifice (cost, time

and trouble) of implementing measures to improve safety with the reduction in risks to those that might suffer harm (HSE, 2014). Thus the amount it is reasonable to spend on safety measures may be judged as a trade-off between the benefit that the system confers in terms of improved safety and the loss of benefit brought about by the costs of paying for those measures. In line with the Kaldor–Hicks compensation principle (Kaldor, 1939; Hicks, 1939), it is customary to assign the cost and hence a notional reduction in wealth to those being protected, even though they will rarely have to meet the bill in

* Corresponding author. Tel.: +44 117 331 5830; fax: +44 117 954 5208.

E-mail addresses: pjt3.michaelmas@gmail.com, philip.thomas@bristol.ac.uk (P.J. Thomas).
<http://dx.doi.org/10.1016/j.psep.2015.07.002>

0957-5820/© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

practice. Utility functions are then used to characterise the fall in benefit these people experience as a result of their assumed reduction in wealth.

To compare the outlay on implementation with the costs in terms of loss of life and other detriments through a cost-benefit analysis (CBA), it is fundamentally necessary to place a monetary value on human life. The “value of a prevented fatality” (VPF) derives from one such valuation exercise. A figure for the VPF is published annually by the UK Government’s [Department for Transport \(2013\)](#) and this is used also by the Health and Safety Executive (HSE) and the Office for Nuclear Regulation (ONR) in developing new safety regulations and determining whether safety measures meet the legal requirements—though it should be emphasised that the latter are the responsibility of those carrying out the work activity.

It is clear that ensuring the VPF is a true reflection of what should be spent is very important and to this end various methods of eliciting the VPF from social and economic surveys have been attempted. The VPF that is currently used in the UK is based on a study by [Carthy et al. \(1999\)](#), which uses a two-injury chained method, whereby individuals are asked to consider two serious injuries, with injury X more severe than injury W. The individual is asked to estimate the maximum acceptable price (MAP) he would pay to avert the specified injury and the minimum acceptable compensation (MAC) he would take as compensation for enduring the injury. The analysis of the data makes very extensive use of utility functions in an attempt to find the amount it is notionally reasonable to pay for a safety measure that will reduce by one the expected number of premature deaths in a large population, given that those deaths are associated with a specified hazard. This sum is deemed to be the VPF.

This approach to deriving a figure for the VPF originated in the transport sector but the concept is now applied much more widely. After tracing the history of the development of the UK VPF figure, [Wolff and Orr \(2009\)](#) concluded that:

“it appears that the Carthy study ([Carthy et al., 1999](#)) is now the primary source of VPF figures, adjusted for inflation and changes in GDP.”

A 2011 report for the Department for Transport (DfT), with authors in common with the Carthy study, recommended “against any early new full scale WTP [willingness to pay] study” ([Spackman et al., 2011](#)). Thus the survey conducted by Carthy et al. of 167 people in 1997 ([Carthy et al., 1999](#)) remains the evidential base for the VPF used by the Government, regulators and many industries in the UK today, including the process and nuclear industries. It is obviously of crucial importance, therefore, that the Carthy study should be soundly based.

Updated for increases in GDP per head, the VPF is assumed to be the same for all people in the UK, irrespective of age or gender. While this might be a dubious assumption (see [Nathwani et al., 1997, 2008](#); [Pandey and Nathwani, 2003](#); [Pandey et al., 2006](#); [Sunstein, 2004a,b](#); [Thomas et al., 2006a, 2006b, 2010a](#), [Thomas and Vaughan, 2013](#)), the VPF is used extensively in the UK.

We have questioned in earlier papers ([Thomas and Vaughan, 2015a,b](#)) the methodology used by Carthy et al, showing that their work contained serious flaws. Using the Carthy authors’ own data it was demonstrated that the method was invalid in that it failed to estimate consistently a key parameter determining the size of the VPF. This parameter, m_{Xi} , is the individual’s marginal rate of substitution

between his wealth and his probability of not suffering injury X. The Carthy study’s data allow estimation of m_{Xi} in two different ways, giving $m_{Xi}^{(1)}$ and $m_{Xi}^{(2)}$ respectively. These values should, of course, be equal for the method to be sound, but they are, in fact, very different and, indeed, barely correlated. Thus the two-injury chained method has been falsified in the sense used by [Popper \(1934\)](#). Indeed, whilst still defending their methodology, the Carthy co-authors have admitted to methodological problems:

“there is a definite and seemingly systematic divergence between direct and indirect estimates which is illustrated by the comparison between $m_{Xi}^{(2)}$ and $m_{Xi}^{(1)}$ ” ([Chilton et al., 2015](#)).

Moreover, [Thomas and Vaughan \(2015b\)](#) found that many of the defences put forward by the Carthy authors in an effort to justify their methodology were flawed or mistaken and so concluded that their attempt to support the use of the methodology was not tenable.

A sub-group of those involved in the first defence published a second attempt at a justification of the Carthy study ([Jones-Lee and Loomes, 2015](#)). While not disputing the failure of the two-injury chained method in the fundamental validity test just mentioned, they raised two main points as a follow on to [Thomas and Vaughan \(2015b\)](#), concerning

- (1) the form the utility function for wealth should take if the individual puts his maximum acceptable price (MAP) for averting the injury as high or higher than the minimum acceptable compensation (MAC) he would countenance to endure the injury;
- (2) whether it is legitimate to deduce a wealth for the respondent from his stated MAP and MAC.

In considering these issues, it should be emphasised at the outset that they are both subsidiary to the previously accepted objection of systematic divergence detailed above, which is sufficient on its own to invalidate the two-injury chained method on which the UK VPF rests. We may conclude immediately, therefore, that the points now raised by Jones-Lee and Loomes are insufficiently important to affect let alone overrule the major criticism we put forward in our first paper. Therefore we reiterate our belief that the VPF derived by Carthy et al. and used for the past 16 years in the UK is not satisfactory for use in making safety decisions as it is not based on a sound analysis.

Nevertheless it must be recognised that the VPF figure used in the UK since 1999 rests squarely on the value produced by the Carthy study. It is thus very important that the attempted second defence of that study by Jones-Lee and Loomes should not be regarded as cover for continuing to use an invalid figure.

Much of the argument in [Jones-Lee and Loomes \(2015\)](#) concerned the use of utility functions and it is the purpose of this paper to investigate how the various utility functions may be used legitimately. Indeed the issues raised by [Jones-Lee and Loomes \(2015\)](#) are of both theoretical and practical interest. It is intriguing that different utility functions from the same Exponential family can give the same result for a key intermediate parameter in the calculation of the VPF when the two-injury chained method is used, namely the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth under injury k . The paper will explain why this is so.

Theoretically interesting results are also produced through considering the proposal of Jones-Lee and Loomes that a

second utility function can be used to determine the individual's utility when his wealth lies below the threshold of validity of a first utility function. It will be shown that this procedure introduces a distorting and unrealistic kink “at the join” between the two utility functions, the avoidance of which requires that the wealth threshold be zero or negligibly small. This rules out the suggestion of Jones-Lee and Loomes that two utility functions can be used, one valid above the wealth threshold and the other below. Setting the threshold at a value higher than zero would mean that the individual's utility of wealth would show a sharp and discontinuous rise at the wealth threshold, characteristic only of the very first increment of wealth. It would suggest that the individual does not start to value his wealth at all until it reaches the threshold, which is an unrealistic model of human behaviour unless the threshold level is negligibly small. Hence a single utility function is needed if all positive wealths are to be covered when the first utility function is one of the three types used in the Carthy study that allow a determination of the individual's wealth, namely the Constrained Power, the Logarithmic and the Negative Inverse.

The use of a single utility function, or, equivalently, fixing the wealth threshold at zero, leads to the very low implied individual wealths associated with the Carthy study discussed in [Thomas and Vaughan \(2015a\)](#), now accepted by [Chilton et al. \(2015\)](#) to be “ridiculously low”. Consistency in the size of the wealth threshold, set by Carthy et al. at zero for the Negative Exponential utility function, constituted the rationale for the assumption of [Thomas and Vaughan \(2015a\)](#) that the zero threshold specified for the Negative Exponential utility function would carry across to the other three utility functions used in the Carthy study: Constrained Power, Logarithmic and Negative Inverse. However the avoidance of the unrealistic kink “at the join” between two utility functions, as just discussed, provides another reason why the wealth threshold must be zero.

The additional inconsistency noted previously with the Carthy study remains unchanged, whereby each utility function gives two different average wealths for the same sample population, one when injury X is considered and a different one for injury W. This is a clear nonsense.

Further interesting results are derived if, purely for the purposes of argument and illustration, it is decided to set aside the finding that the wealth threshold for each person should be zero or negligibly small if a utility function is to be used in characterising decisions on safety measures. The two, inconsistent and low implied average wealths found previously turn out to be entirely unchanged when people's behaviour is modelled using the Constrained Power utility function. It is, moreover, possible to use random variable analysis to deduce plausible average wealths under the Logarithmic and Negative Inverse utility functions. Low average wealths are produced once more, and the infeasible split is exacerbated between the two average wealths deduced for the same set of people under the two injuries, X and W, considered in the Carthy study.

The intention of this paper is to lay out a clear explanation of the issues raised by Jones-Lee and Loomes before refuting them in the hope that the results may be of interest and value to the wider safety community. Section 2 will put the use of utility functions in safety analysis into context and then explain the technical issues raised by Jones-Lee and Loomes.

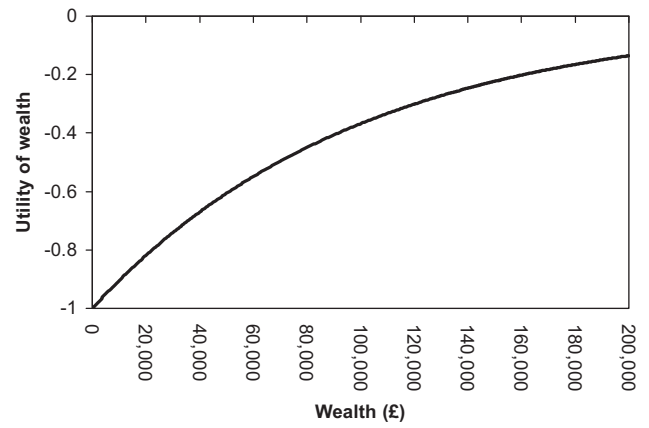


Fig. 1 – Example of a utility function: (Negative–Negative utility function, $U_i = -\exp(-\beta_i w_i)$; $\beta_i = 10^{-5}$).

2. The use of utility functions in safety decision making

2.1. General properties of utility functions

Utility functions embody the concept that a person will value successive increments of his wealth differently, with the first increment normally worth more to him than later increments. [Fig. 1](#) shows as an example a Negative Exponential utility function with parameter, β_i , set to 10^{-5} for illustrative purposes (see Eq. (3) below for the definition of the curve). The fact that the person's utility of wealth is negative is held not to have significance because in utility theory it is utility differences that matter rather than the absolute value of utility. The important thing to observe is that the worth the person sets on his wealth, that is to say his utility, rises as his wealth increases, but not in a linear fashion. He values successive increments less.

A calculation of the appropriate amount to spend on a safety measure to protect human life will often rely on the use of utility functions. These can be applied to characterise the fall in benefit experienced by the people being protected by a safety system as a result of their wealth being notionally reduced to the level it would have been if the burden of payment had actually fallen on them, in line with the Kaldor–Hicks compensation principle.

Utility functions are normally expected to be differentiable, which facilitates analysis, and it is argued in [Thomas \(2010\)](#) that of the differentiable functions, only the Power family of utility functions can offer a realistic model of human decision making. This is because only the Power family stipulates that the person's risk-aversion stays constant during the process of comparing each pair of alternative wealth outcomes. Indeed the Power utility function is widely used by economists and actuaries ([Boadway and Bruce, 1984](#)), and has the basic form:

$$U_i(w_i) = w_i^{1-\varepsilon_i} \quad (1)$$

where $U_i(\cdot)$ is the utility function, w_i is the wealth of individual, i , while ε_i is his dimensionless risk-aversion. This is defined for any type of utility function as:

$$\varepsilon_i = -w_i \frac{d^2 U_i}{dw_i^2} \div \frac{dU_i}{dw_i} \quad (2)$$

For example, the Logarithmic utility function recommended by the U.K. [Treasury \(2011\)](#) is a limiting form of the Power utility function where the risk-aversion is unity: $\varepsilon_i = 1.0$.

It is normal to assume that a positive risk-aversion, indicating caution, will be in force when decisions of consequence are being taken, e.g., when human life is at risk. This implies a utility function that is increasing in wealth but concave—showing a tendency to level out (although there may be no asymptotic top level). Higher risk-aversions will normally result in more cautious decisions (although there are limits; see [Thomas et al., 2010a,b](#); [Thomas and Jones, 2010c](#); [Thomas, 2013](#) for a discussion of these).

Other differentiable functions of roughly the same shape are sometimes suggested as serviceable for representing utility. One such is the Negative Exponential utility function illustrated in [Fig. 1](#):

$$U_i(w_i) = -e^{-\beta_i w_i} \quad (3)$$

where β_i is a constant that is characteristic of the individual, with $\beta_i > 0$ being the canonical form. This produces a utility that is always negative, but increases in a concave fashion from -1.0 to 0.0 as wealth, w_i , increases from zero.

The risk-aversion for individual, i , operating to this utility function is:

$$\varepsilon_i = \beta_i w_i \quad (4)$$

which, for $\beta_i > 0$, is positive for any positive wealth, w_i , indicating a cautious decision-maker.

The Positive Exponential was discussed as a utility function by [Jones-Lee and Loomes \(2015\)](#):

$$U_i(w_i) = e^{\beta_i w_i} \quad (5)$$

with $\beta_i > 0$ producing the canonical form. This returns a value of 1 at $w_i = 0$ and increases exponentially and without limit as wealth goes up. The risk-aversion for individual, i , is now:

$$\varepsilon_i = -\beta_i w_i \quad (6)$$

which, for $\beta_i > 0$, will be negative for any positive wealth, indicating a risk-seeking decision maker.

The reversal of the sign for risk-aversion suggests that the Positive Exponential utility function, as defined by Eq. (5), can be regarded in a sense as the opposite of the Negative Exponential utility function defined by Eq. (3).

2.2. The effect of MAP being equal or higher than MAC on the utility functions used in the 1999 study

[Thomas and Vaughan \(2015a\)](#) pointed out that all 4 utility functions used by the Carthy study suffered invalidating problems when $\text{MAP} \leq \text{MAC}$. (Besides the Negative Exponential utility function, the Carthy study tabulated results from the Constrained Power, Logarithmic and Negative Inverse utility functions. These utility functions are defined and discussed further in Section 7 below.) Nevertheless it was found that a Negative Exponential utility function, when pushed beyond the limits of its applicability, could reproduce the common results that the Carthy authors decided to write across all utility functions for the 30% of cases where $\text{MAP} \leq \text{MAC}$. These cases are listed in the document, “Individual Responses & VOSLS-PEG Study”, HSE/Peg/Nov 1997/SC, supplied by

Table 1 – Pairs of Exponential utility functions that apply for ranges of MAC to MAP ratio, $c_{ki} = y_{ki}/x_{ki}$.

	$1 < c_{ki} < \infty$	$c_{ki} = 1$	$0 < c_{ki} < 1$
Negative Exponential	NNEx	NZEx	NPEX
Positive Exponential	PNEx	PZEx	PPEX

Professor Jones-Lee, which contains the data used in the Carthy study.

The two-injury chained method used in the 1999 study implies that the value of β_i may vary, leading in some cases to $\beta_i = 0$ and in others to $\beta_i < 0$, as well as the more normal $\beta_i > 0$. Some interesting and important links between the behaviour of the two basic Exponential utility functions then emerge, as indicated by [Jones-Lee and Loomes \(2015\)](#).

In their response to [Thomas and Vaughan \(2015a\)](#), [Chilton et al. \(2015\)](#) complained that:

“the Negative Exponential utility function produces a completely different result from the one that we derived and reported in the data that we sent to them.” [their italics]

but this statement was refuted comprehensively in [Table 1](#) of [Thomas and Vaughan \(2015b\)](#). However, [Jones-Lee and Loomes \(2015\)](#) have now clarified that they themselves were using a Positive Exponential utility function, of the form $U_i(w_i) = e^{\beta_i w_i}$ with $\beta_i > 0$, to cover cases where $\text{MAP} < \text{MAC}$ and a Linear utility function, of the form $U_i(w_i) = g_i w_i + h_i$ with $g_i > 0$, when $\text{MAP} = \text{MAC}$.

On this basis, [Jones-Lee and Loomes \(2015\)](#) describe the use of the Negative Exponential to derive their results as a “fundamental error”. In fact, as Jones-Lee and Loomes admit, the Positive Exponential and Linear utility functions produce exactly the same answers as found using the Negative Exponential utility function when operated outside its normal range. It is thus objectively impossible to divine from the tabulated results which of the alternative utility functions were used.

Nevertheless the point is of interest. Both the Positive Exponential and the Negative Exponential utility function belong to the family of Exponential utility functions, and this paper will explore in Sections 3–6 how different utility functions in the same Exponential family can give the same result for a key intermediate parameter in the calculation of the VPF using the two-injury chained method. This parameter is the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth.

[Jones-Lee and Loomes \(2015\)](#) have now further made it clear that the Carthy study was advocating and employing a switch between different utility functions when $\text{MAP} \geq \text{MAC}$, even if the labels in the tables of results in the Carthy study were left bearing the name of the default utility function applied to the rest of the cohort for whom $\text{MAP} < \text{MAC}$. By contrast [Thomas and Vaughan \(2015a\)](#) had achieved the same figures by assuming a movement between different modes of the Negative Exponential utility function, thus ensuring continuity of utility at least in the case of the Negative Exponential utility function.

Thomas and Vaughan were, however, puzzled as to the justification for writing the results produced by the Negative Exponential utility function when pushed beyond its normal limits in place of results from the Constrained Power, the Logarithmic and the Negative Inverse utility functions. This raises the question as to whether it is acceptable to

make the change to the Positive Exponential utility function from even its closest relative, the Negative Exponential utility function. The paper will show such a leap to be a dubious proposition.

2.3. Deducing the wealth of respondents

2.3.1. The status of the wealth threshold

Jones-Lee and Loomes (2015) advocate a further switch between utility functions, this time between the nominal utility function, taken to be valid above individual i 's wealth threshold, $\beta^{(i)} \geq 0$, and a second, auxiliary utility function to cover wealths less than $\beta^{(i)}$. (The wealth threshold, $\beta^{(i)}$, is not related to the parameter, β_i , used to characterise the Exponential utility function, but the notation follows that used in the Carthy study.)

Jones-Lee and Loomes suggest that the second utility function can be fitted underneath the first. Invoking this auxiliary utility function might then get around the fact, pointed out in Thomas and Vaughan (2015b), that any strictly positive wealth threshold would leave the individual's utility undefined for wealths below $\beta^{(i)}$. If the argument could be sustained that a separate utility function can be invoked to cover individual i 's wealth, w_i , when it falls below the threshold, $w_i < \beta^{(i)}$, then what Jones-Lee and Loomes describe as the “ridiculously low implied levels of wealth” that are an embarrassment for the two-injury chained method would be transformed into wealth differences only. Discrepancies amongst wealth differences might then be considered less damaging to the validity of the two-injury chained method.

The paper will therefore examine (in Section 7) the proposal that a second utility function valid for wealths up to the wealth threshold, $\beta^{(i)}$, may be joined onto the first utility function that covers the higher wealths. It will be shown that the proposition is untenable. The inclusion of a wealth threshold in the utility function precludes a meaningful definition of utility for any wealth less than the threshold, a situation that must be unrealistic unless $\beta^{(i)} \rightarrow 0$.

2.3.2. Examination of the situation when a non-zero wealth threshold is assumed to be possible

In Section 8, it will be assumed, for the sake of argument, that it is possible to include an incompletely defined wealth threshold, $\beta^{(i)}$, in the definition of the utility function, as advocated by Jones-Lee and Loomes (2015).

It will be shown that, even in this case, the hope of avoiding the “ridiculously low implied levels of wealth” turns out to be vain. It will be shown that it is still possible to deduce the wealth of each individual in the case of the Constrained Power utility, and hence the average wealth in the cohort for comparison with the UK average wealth at the time. Moreover the values are unaltered from those reported in Thomas and Vaughan (2015a), about 5% of the wealth of the average UK adult at the time of the Carthy study.

In the cases of the Logarithmic and Negative Inverse utility functions, it is possible, by employing a random variable analysis, to make a reasonable estimate of the expected average wealth of the cohort, a figure suitable once again for comparison processes. The resultant low values and discrepancies demonstrate the invalidity of the two-injury chained model even if a case for retaining the wealth threshold could be sustained (which is not possible).

3. The use of the Negative Exponential utility function in the two-injury chained method

3.1. Relating the MAC to MAP ratio, c_{ki} , to the parameter, β_i , that defines the shape of the Negative Exponential utility function

Appendix A relates the individual's starting utility of wealth to his utilities after paying the maximum acceptable price (MAP), x_{ki} , to avert the injury and after receiving his minimum acceptable compensation (MAC), y_{ki} , for enduring the injury. The following equation is derived:

$$U_i(w_{i0} - x_{ki}) + U_i(w_{i0} + y_{ki}) = 2U_i(w_{i0}) \quad k = W, X \quad (A.8)$$

where w_{i0} is the individual's starting wealth.

Substituting from Eq. (3) into Eq. (A.8) gives an expression relevant to the Negative Exponential utility function:

$$e^{-\beta_i w_{i0}} e^{\beta_i x_{ki}} + e^{-\beta_i w_{i0}} e^{-\beta_i y_{ki}} = 2e^{-\beta_i w_{i0}} \quad (7)$$

Multiplying throughout by $e^{\beta_i w_{i0}}$ gives

$$e^{\beta_i x_{ki}} + e^{-\beta_i y_{ki}} = 2 \quad (8)$$

which implies that all further results will apply irrespective of the individual's starting wealth, w_{i0} . Now define c_{ki} as the individual's MAC to MAP ratio:

$$c_{ki} = \frac{y_{ki}}{x_{ki}} \quad (9)$$

where $c_{ki} \geq 0$. Eq. (8) now becomes:

$$e^{\beta_i x_{ki}} + e^{-\beta_i c_{ki} x_{ki}} = e^{\beta_i x_{ki}} + (e^{-\beta_i x_{ki}})^{c_{ki}} = 2 \quad (10)$$

Using the further substitution,

$$\phi_{ki} = e^{\beta_i x_{ki}} \quad (11)$$

yields:

$$\phi_{ki} - 2 + \frac{1}{\phi_{ki}^{c_{ki}}} = g(\phi_{ki}, c_{ki}) = 0 \quad (12)$$

where, the function, $g(\phi_{ki}, c_{ki})$, is as defined above. It may be noted that $c_{ki} = 1$ turns Eq. (12) into the quadratic equation, $\phi_{ki}^2 - 2\phi_{ki} + 1 = 0$, which has $\phi_{ki} = 1$ as its sole solution. This knowledge facilitates an iterative solution for ϕ_{ki} at two values of c_{ki} , one just below unity and the other marginally above. These solutions provide starting points for the Method of Referred Derivatives (Thomas, 1999), which takes advantage of the analytic continuity of the function, $g(\phi_{ki}, c_{ki})$, to find ϕ_{ki} in terms of c_{ki} over the range $0 < c_{ki} < \infty$. See Fig. 2, which shows a smooth curve passing through the point (1, 1) with asymptotic values $\phi_{ki} \rightarrow 0$ as $c_{ki} \rightarrow 0$ and $\phi_{ki} \rightarrow 2$ as $c_{ki} \rightarrow \infty$. Thus we may write

$$\phi_{ki} = F(c_{ki}) = F\left(\frac{y_{ki}}{x_{ki}}\right) \quad (13)$$

where, $F(y_{ki}/x_{ki})$ is given graphically in Fig. 2.

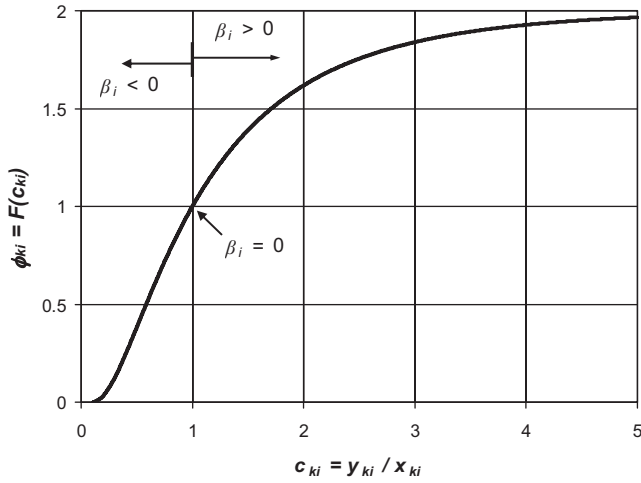


Fig. 2 – Plot of ϕ_{ki} versus c_{ki} with the regions of β_i marked up: positive, negative and zero.

It follows from $\phi_{ki} = e^{\beta_i x_{ki}}$ that

$$\beta_i = \frac{\ln \phi_{ki}}{x_{ki}} \quad (14)$$

Since the MAP, x_{ki} , can be expected to be greater than zero, β_i will take the sign of $\ln \phi_{ki}$. The curve of ϕ_{ki} vs. c_{ki} shown in Fig. 2, taken in conjunction with Eq. (14), implies the following conditions:

$$\begin{aligned} 0 < c_{ki} < 1 &\Rightarrow \phi_{ki} < 1 \Rightarrow \beta_i < 0 \\ c_{ki} = 1 &\Rightarrow \phi_{ki} = 1 \Rightarrow \beta_i = 0 \\ 1 < c_{ki} < \infty &\Rightarrow \phi_{ki} > 1 \Rightarrow \beta_i > 0 \end{aligned} \quad (15)$$

The conditions on β_i have been marked up on Fig. 2.

3.2. The marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth, w_i , under the two-injury, chained method

Substituting from Eq. (3) into Eq. (A.9) gives the injury offset, a_{ki} , for individual, i , and injury, k , for the Negative Exponential utility function as:

$$\begin{aligned} a_{ki} &= -e^{-\beta_i w_{i0}} + e^{-\beta_i (w_{i0} - x_{ki})} = e^{-\beta_i w_{i0}} (e^{\beta_i x_{ki}} - 1) \\ &= -U_i(w_{i0}) (\phi_{ki} - 1) \end{aligned} \quad (16)$$

Meanwhile, differentiating Eq. (3) with respect to w_i gives, at $w_i = w_{i0}$:

$$\frac{dU_i}{dw_i} = \beta_i e^{-\beta_i w_{i0}} = -\frac{\ln \phi_{ki}}{x_{ki}} U_i(w_{i0}) \quad (17)$$

where, Eqs. (3) and (14) have been used in the second step.

The marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth, w_i , is given by Eq. (A.10), introduced in Appendix A:

$$m_{ki} = \frac{a_{ki}}{\frac{dU_i}{dw_i}} \quad k = W, X \quad (A.10)$$

Hence, substituting from Eqs. (16) and (17) into Eq. (A.10) gives the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , as

$$m_{ki} = \frac{\phi_{ki} - 1}{\ln \phi_{ki}} x_{ki} \quad (18)$$

with k taking the value, W or X , depending on which injury is under consideration. Using Eq. (13), we may express this as:

$$m_{ki} = \frac{F\left(\frac{y_{ki}}{x_{ki}}\right) - 1}{\ln F\left(\frac{y_{ki}}{x_{ki}}\right)} x_{ki} \quad k = W, X \quad (19)$$

3.3. Properties of the 3 categories of the Negative Exponential utility function

As shown in condition set (15) and as marked on Fig. 2, three regions may be defined for β_i where, $\beta_i < 0$, $\beta_i = 0$ and $\beta_i > 0$. These demarcate 3 different categories of the Negative Exponential utility function. The first category, where $\beta_i > 0$, embodies two negatives: the minus sign immediately after the equals sign in Eq. (3) and the negative exponent. This, the canonical form, may be named the Negative-Negative Exponential (NNEx) utility function in consequence.

The second form, with $\beta_i = 0$, retains the minus sign immediately after the equals sign in Eq. (3) but has a zero exponent and so may be characterised as the Negative-Zero Exponential (NZEx) utility function. The third form, when $\beta_i < 0$, keeps the minus sign immediately after the equals sign in Eq. (3) but has a positive exponent. This form will be called the Negative-Positive Exponential (NPEx) utility function.

3.3.1. The Negative-Negative Exponential (NNEx) utility function ($\beta_i > 0$)

From condition set (15), values of c_{ki} in the range $1 < c_{ki} < \infty$ will lead to $\beta_i > 0$, leading to NNEx utility function, given by

$$U_i(w_i) = -e^{-\beta_i w_i} \quad \beta_i > 0 \quad (20)$$

which produces a utility that is always negative, but increases in a concave fashion from -1 to 0 as wealth, w_i , increases from 0 . The risk-aversion, ε_i , for individual, i , is positive, indicating caution, a caution that increases with wealth:

$$\varepsilon_i = \beta_i w_i \quad \beta_i > 0 \quad (21)$$

3.3.2. The Negative-Positive Exponential (NPEx) utility function ($\beta_i < 0$)

From condition set (15), values of c_{ki} in the range $0 < c_{ki} < 1.0$ will lead to $\beta_i < 0$ and so to the NPEx utility function, defined by

$$U_i(w_i) = -e^{-\beta_i w_i} \quad \beta_i < 0 \quad (22)$$

This might be regarded as an anti-materialist utility function, since it models a person who prefers less wealth to more. (We shall suspend judgement at this stage on whether or not this is a realistic model of human choice concerning safety measures.)

The risk-aversion for individual, i , is negative indicating that the person is risk-seeking, a tendency that increases with wealth:

$$\varepsilon_i = \beta_i w_i \quad \beta_i < 0 \quad (23)$$

3.3.3. Properties of the Negative–Zero Exponential (NZEx) utility function ($\beta_i = 0$)

From condition set (15), a unity value of c_{ki} : $c_{ki} = 1$ will lead to $\beta_i = 0$, and hence the NZEx utility function, which is given by:

$$U_i(w_i) = -e^{-\beta_i w_i} \big|_{\beta_i=0} = -1 \quad (24)$$

As indicated by Fig. 2, it is the limiting form of both the NNEx utility function and the NPEx utility function as $\beta_i \rightarrow 0$ and $\phi_{ki} \rightarrow 1$. It models the behaviour of someone who is indifferent to wealth.

From Eq. (4), the risk-aversion is zero, a feature that normally indicates a risk-neutral individual:

$$\varepsilon_i = 0 \quad (25)$$

Applying l'Hôpital's rule to Eq. (18), shows that the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , obeys

$$m_{ki} \rightarrow x_{ki} \text{ as } c_{ki} \rightarrow 1 \quad (26)$$

where $\phi_{ki} \rightarrow 1$ as $c_{ki} \rightarrow 1$.

3.4. Summary of the properties of Negative Exponential utility functions

The Negative–Negative Exponential (NNEx) utility function is associated with a MAC to MAP ratio that lies in the range: $1 < c_{ki} < \infty$. This leads to a strictly positive value for the beta parameter governing the shape of the utility function: $\beta_i > 0$. The utility function is increasing but concave and the risk-aversion is positive, indicating cautious behaviour.

The Negative–Positive Exponential (NPEx) utility function is produced when the MAC to MAP ratio lies in the range: $0 < c_{ki} < 1$. This leads to the beta parameter being strictly negative, $\beta_i < 0$, causing the utility function to be decreasing but concave. The risk-aversion is negative, a property that normally indicates risk-seeking behaviour.

The Negative–Zero Exponential (NZEx) utility function results from a MAC to MAP ratio of unity: $c_{ki} = 1$. This leads to a zero value for the beta parameter: $\beta_i = 0$. The utility function is then a horizontal straight line. The risk-aversion is zero, something that normally indicates risk-neutrality.

Discussion of the validity of the NPEx and NZEx utility functions as models for human behaviour will be deferred until the end of Section 6.

4. The use of the Positive Exponential utility function in the two-injury chained method

4.1. Relating the MAC to MAP ratio, c_{ki} , to the parameter, β_i , defining the shape of the Positive Exponential utility function

Substituting from Eq. (5) into Eq. (A.8) gives

$$e^{\beta_i w_{i0}} e^{-\beta_i x_{ki}} + e^{\beta_i w_{i0}} e^{\beta_i y_{ki}} = 2e^{\beta_i w_{i0}} \quad (27)$$

which may be reduced to

$$e^{-\beta_i x_{ki}} + e^{\beta_i y_{ki}} = 2 \quad (28)$$

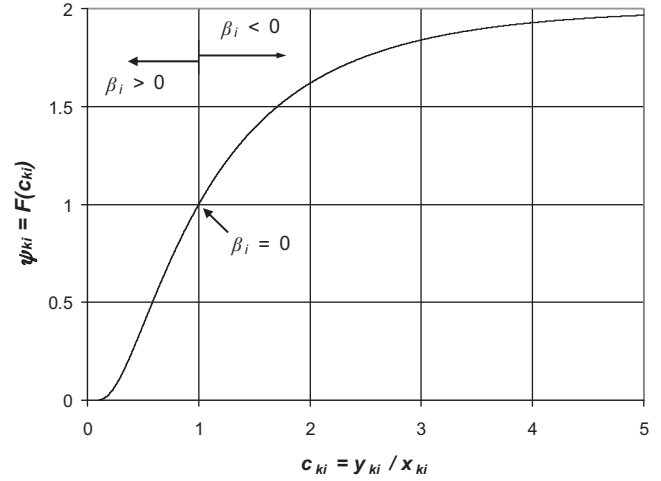


Fig. 3 – Plot of ψ_{ki} vs. c_{ki} with the regions of β_i marked up: positive, negative and zero.

by multiplying throughout by $e^{-\beta_i w_{i0}}$. Substituting from Eq. (9), namely $c_{ki} = y_{ki}/x_{ki}$, into Eq. (27) gives:

$$e^{-\beta_i x_{ki}} + e^{\beta_i c_{ki} x_{ki}} = e^{-\beta_i x_{ki}} + (e^{\beta_i x_{ki}})^{c_{ki}} = 2 \quad (29)$$

Applying the further substitution,

$$\psi_{ki} = e^{-\beta_i x_{ki}} \quad (30)$$

yields:

$$\psi_{ki} - 2 + \frac{1}{\psi_{ki}^{c_{ki}}} = 0 \quad (31)$$

Eq. (31) is of the same form as Eq. (12), and so its solutions will reproduce those shown in Fig. 2. Hence

$$\psi_{ki} = F(c_{ki}) = F\left(\frac{y_{ki}}{x_{ki}}\right) \quad (32)$$

However, since now, from Eq. (30)

$$\beta_i = -\frac{\ln \psi_{ki}}{x_{ki}} \quad (33)$$

the regions where $\beta_i < 0$ and $\beta_i > 0$ will be reversed. See Fig. 3. Hence

$$\begin{aligned} 0 < c_{ki} < 1 &\Rightarrow \psi_{ki} < 1 \Rightarrow \beta_i > 0 \\ c_{ki} = 1 &\Rightarrow \psi_{ki} = 1 \Rightarrow \beta_i = 0 \\ 1 < c_{ki} < \infty &\Rightarrow \psi_{ki} > 1 \Rightarrow \beta_i < 0 \end{aligned} \quad (34)$$

Substituting from Eq. (5) into Eq. (A.9) gives the injury offset, a_{ki} , for individual, i , and injury, k , for the Positive Exponential utility function as:

$$\begin{aligned} a_{ki} &= e^{\beta_i w_{i0}} - e^{\beta_i (w_{i0} - x_{ki})} = e^{\beta_i w_{i0}} (1 - e^{-\beta_i x_{ki}}) \\ &= -U_i(w_{i0}) (\psi_{ki} - 1) \end{aligned} \quad (35)$$

Meanwhile, differentiating Eq. (5) with respect to w_i gives, at $w_i = w_{i0}$:

$$\frac{dU_i}{dw_i} = \beta_i e^{\beta_i w_{i0}} = -\frac{\ln \psi_{ki}}{x_{ki}} U_i(w_{i0}) \quad (36)$$

Since $\psi_{ki} = \exp(-\beta_i x_{ki})$ implies $\beta_i = -\ln \psi_{ki}/x_{ki}$ it follows that substituting from Eqs. (35) and (36) into Eq. (A.10) gives the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , as

$$m_{ki} = \frac{\psi_{ki} - 1}{\ln \psi_{ki}} x_{ki} \quad (37)$$

where, k may take the value W or X . Using Eq. (32) we may write

$$m_{ki} = \frac{F\left(\frac{y_{ki}}{x_{ki}}\right) - 1}{\ln F\left(\frac{y_{ki}}{x_{ki}}\right)} x_{ki} \quad k = W, X \quad (38)$$

Comparing Eqs. (19) and (38), it is clear that the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , will always be the same for the Positive Exponential utility function as for the Negative Exponential utility function for any combination of MAP and MAC, (x_{ki}, y_{ki}) .

4.2. Properties of the 3 categories of the Positive Exponential utility function

Allowing for the 3 cases where, $\beta_i > 0$, $\beta_i = 0$ and $\beta_i < 0$, the Positive Exponential utility function may also be split into 3 separate categories. The first, canonical form, where, $\beta_i > 0$, embodies two positives: the implied plus sign immediately after the equals sign in Eq. (5) and the positive exponent. This form will be specified as the Positive–Positive Exponential (PPEX) utility function.

The second form, with $\beta_i = 0$, retains the implied plus sign immediately after the equals sign in Eq. (5) but has a zero exponent. This will be characterised as the Positive–Zero Exponential (PZEX) utility function. The third form, when $\beta_i < 0$, keeps the implied plus sign immediately after the equals sign in Eq. (5) but has a negative exponent. This form will be called the Positive–Negative Exponential (PNEX) utility function.

4.2.1. The Positive–Positive Exponential (PPEX) utility function ($\beta_i > 0$)

From condition set (34), values of c_{ki} in the range $0 < c_{ki} < 1.0$ mean that $\beta_i > 0$, which leads to the PPEX utility function, which has the form:

$$U_i(w_i) = e^{\beta_i w_i} \quad \beta_i > 0 \quad (39)$$

This produces a utility that is always positive, but increases in a convex fashion from $(0, 1)$ with wealth, w_i . The risk-aversion for individual, i , is negative and decreasing with wealth, indicating he will become ever more risk-seeking as he gets richer:

$$\varepsilon_i = -\beta_i w_i \quad \beta_i > 0 \quad (40)$$

4.2.2. The Positive–Negative Exponential (PNEX) utility function ($\beta_i < 0$)

From condition set (34), values of c_{ki} in the range $1 < c_{ki} < \infty$ mean that $\beta_i < 0$, which leads to the PNEX utility function, which has the form:

$$U_i(w_i) = e^{\beta_i w_i} \quad \beta_i < 0 \quad (41)$$

Like the NPEx utility function, this may be regarded as an anti-materialist utility function, since it returns a unity

utility at zero wealth and then declines, thus modelling a person who prefers less wealth to more. We shall once again suspend judgement on the realism of this utility function until later.

The risk-aversion for individual, i , is positive and increasing with wealth, indicating an increasingly risk-averse person as he gets wealthier:

$$\varepsilon_i = -\beta_i w_i \quad \beta_i < 0 \quad (42)$$

4.2.3. The Positive–Zero Exponential (PZEX) utility function ($\beta_i = 0$)

From condition set (34), a unity value for c_{ki} : $c_{ki} = 1.0$ means that $\beta_i = 0$, which implies the PNEX utility function, with the form:

$$U_i(w_i) = e^{\beta_i w_i} |_{\beta_i=0} = 1 \quad (43)$$

As indicated by Fig. 3, it is the limiting form of both the PPEX utility function and the PNEX utility function as $\psi_{ki} \rightarrow 1$ and so $\beta_i \rightarrow 0$. Like the NZEX utility function, the PZEX models the behaviour of someone who is indifferent to wealth. The risk-aversion for individual, i , is zero, indicating a risk-neutral person:

$$\varepsilon_i = 0 \quad (44)$$

Applying l'Hôpital's rule to Eq. (37) shows that the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , obeys

$$m_{ki} \rightarrow x_{ki} \quad \text{as} \quad c_{ki} \rightarrow 1 \quad (45)$$

where, $\psi_{ki} \rightarrow 1$ as $c_{ki} \rightarrow 1$.

4.3. Summary of the properties of Positive Exponential utility functions

The Positive–Positive Exponential (PPEX) utility function is associated with a MAC to MAP ratio that lies in the range: $0 < c_{ki} < 1$. This leads to a strictly positive value for the beta parameter governing the shape of the utility function: $\beta_i > 0$. The utility function is convex and increasing. The risk-aversion is negative, indicating risk-seeking behaviour.

The Positive–Negative Exponential (PNEX) utility function is produced when the MAC to MAP ratio lies in the range: $1 < c_{ki} < \infty$. This leads to the beta parameter being strictly negative, $\beta_i < 0$, causing the utility function to be convex but decreasing. The risk-aversion is positive, which normally indicates cautious behaviour.

The Positive–Zero Exponential (PZEX) utility function results from a MAC to MAP ratio of unity: $c_{ki} = 1$. This leads to a zero value for the beta parameter: $\beta_i = 0$. The utility function is then a horizontal straight line. The risk-aversion is zero, which normally indicates risk-neutrality.

The ultimate validity or otherwise of the Positive Exponential utility functions as models for human decisions on safety will be discussed at the end of in Section 6.

5. Similarities and differences between the various Exponential utility functions

Fig. 4 illustrates two groups of utility functions: the three Negative Exponential utility functions, namely NNEX, NZEX and

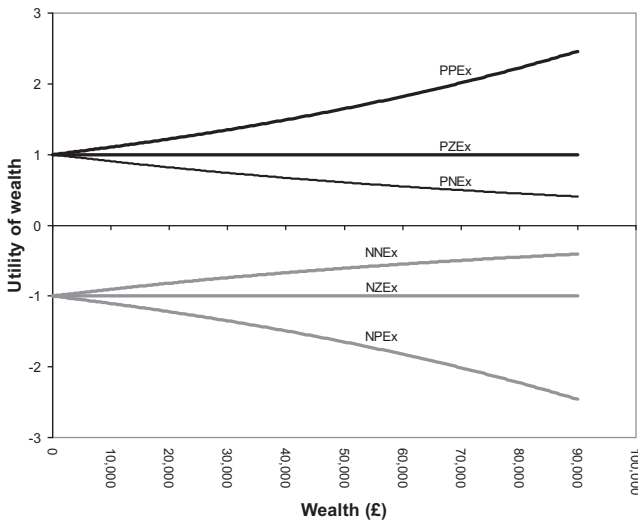


Fig. 4 – Exponential utility functions with $|\beta_i| = 1 \times 10^{-5}$
Upper half: Positive–Positive, Positive–Zero,
Positive–Negative. Lower half: Negative–Negative,
Negative–Zero, Negative–Positive.

NPEx, and the three Positive Exponential utility functions: PPEx, PZEx and PNEx. The defining parameter, β_i , conforms in all cases to

$$|\beta_i| = 1 \times 10^{-5} \quad (46)$$

a value chosen for illustrative purposes. It is clear that the one group is the mirror image of the other.

As pointed out in Section 4.1, a comparison of Eqs. (19) and (38) shows that the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , will always be the same for the Positive Exponential utility function as for the Negative Exponential utility function for any combination of MAP and MAC, (x_{ki}, y_{ki}) . This duality lies behind many of the comments contained in Jones-Lee and Loomes (2015).

Interestingly, despite those authors inveighing there against the calculational framework for m_{ki} developed for a Negative Exponential utility function being employed to calculate m_{ki} for a Positive Exponential utility function, it was suggested in the Carthy study that just such a route might be used:

“with the positive exponential form, for any non-fatal injury I, m_i is given by simply letting \hat{x} and \hat{y} “swap places” in the formulae for the negative exponential case”

In fact it is shown in Appendix C to this paper that it is indeed possible to compute m_{ki} under the Positive Exponential utility function by replacing y_{ki} by x_{ki} and x_{ki} by y_{ki} in the formula for m_{ki} under the Negative Exponential, but there is no need to make this transposition, as the unmodified formula will give the same result. This interesting outcome is a necessary consequence of the model used in the two-injury chained method.

The value of the MAC to MAP ratio, $c_{ki} = y_{ki}/x_{ki}$, determines which of the three Negative Exponential utility functions applies and which of the three Positive Exponential utility functions applies. Table 1 shows the relationship between category of Exponential utility function and MAC to MAP ratio, c_{ki} .

The question now arises does either the NPEx or the PPEx utility function have greater validity than the other? While the

risk-seeking nature of both militates against a convincing case for validity in the context of safety decisions, it can be argued that the NPEx utility function has a stronger claim by virtue of being closer to the NNEx utility function, taken to be one of the default utility functions used in the Carthy study. The nature of this closeness will be discussed in the next section.

6. Continuity of utility: modelling the individual's behaviour through moving from one Exponential utility function to another utility function selected from the Exponential family

Assume the decision has been taken to model the behaviour of a cohort of respondents using a utility function from the Negative Exponential group. Referring to the bottom half of Fig. 4, suppose that someone's MAC, y_{ki} , is slightly greater than his MAP, x_{ki} . His decision-making is thus currently modelled, reasonably, using a NNEx utility function, which we may regard as the default utility function. Now imagine that this individual is considering edging his MAC down so that it will be slightly lower than his MAP. On the way, he considers the possibility that his MAC and MAP should be equal: $y_{ki} = x_{ki}$, implying $c_{ki} = 1$. Modelling this situation using the NZEx utility function will mean that a seamless transfer from the NNEx utility function will have taken place, and this will give him a slightly lower utility of wealth as a result.

Now assume he lowers his MAC slightly more, so that it lies a little below his MAP. Now $y_{ki} < x_{ki}$ and so $c_{ki} < 1$. This will transfer him to the NPEx utility function, and will give him a utility of wealth that is slightly lower again.

The whole process, NNEx \rightarrow NZEx \rightarrow NPEx, will result in a gradual and continuous decline in his utility of wealth.

[Movements among $y_{ki} > x_{ki}$, $y_{ki} = x_{ki}$ and $y_{ki} < x_{ki}$ leading to $c_{ki} > 1$, $c_{ki} = 1$ and $c_{ki} < 1$ respectively, cannot be considered improbable in view of the evidence that 50 out of the 167 respondents in the Carthy survey set $c_{ki} \leq 1$ when considering either injury W or the more severe injury X. Of these 50 people, nearly half, 23, chose a different option for c_{ki} under injury W from the one they chose under injury X.]

Examine now what happens when utility functions taken from the top half as well as the bottom half of Fig. 4 are applied, viz selections may be made from the Positive Exponential group as well as the Negative Exponential group. Let the person's starting behaviour, with $y_{ki} > x_{ki}$, be modelled fairly by a NNEx utility function as previously. Moreover, when he moves to $y_{ki} = x_{ki}$, let his behaviour be modelled, as before, by a NZEx utility function, so that his utility of wealth drops slightly. Now let him make his MAC slightly smaller than his MAP, so that now $y_{ki} < x_{ki}$ and $c_{ki} = (y_{ki}/x_{ki}) < 1$. If his behaviour is now modelled by a PPEx utility function, his utility of wealth will shoot up in a dramatic fashion, analogously to finding a ladder in the snakes-and-ladders board game. For example, if he possesses wealth of £70,000, he will suddenly find that he has more utility under the PPEx than £70 bn would have afforded him under his original, NNEx utility function. The process may be represented as NNEx \rightarrow NZEx \uparrow PPEx, and is clearly highly discontinuous.

His other route to such an unlikely situation under the PPEx utility function is to move seamlessly from the NNEx utility function to the NZEx utility function when $y_{ki} = x_{ki}$, and then transfer effortlessly but discontinuously to the PPEx utility function, bringing a huge gain in utility once more: a step from -1.0 to $+1.0$. Only after this discontinuity can a smooth,

subsequent move be effected into the region of the PPEX utility function, where $y_{ki} < x_{ki}$. The process may be represented as $NNEx \rightarrow NZEx \uparrow PZEx \rightarrow PPEX$ and is again highly discontinuous.

Whichever route is chosen, moving from the NNEx utility function to the PPEX utility function must involve a discontinuous leap in utility.

So considering the case where the default utility function is the NNEx, even though the NPEX utility function and the PPEX utility function give the same marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , the NPEX utility function has a better claim to validity than the PPEX utility function, since it involves only incremental changes in utility in the transfer from the region governed by the NNEx utility function to that where the NPEX utility function operates. By contrast moving to the PPEX utility function must bring a discontinuous leap in utility.

A similar argument holds for the suggestion that a Linear utility function should be used for results presented under the heading of the NNEx utility function when $MAC = MAP$: $y_{ki} = x_{ki} \Rightarrow c_{ki} = 1$. The Linear utility function is a limiting case of the Power utility function when risk-aversion, ε_i , is set to zero in Eq. (1). Allowing for a positive linear transformation, it may be written

$$U_i(w_i) = g_i w_i + h_i \quad g_i > 0 \quad (47)$$

Assuming that the person's starting behaviour, with $y_{ki} > x_{ki}$, is modelled reasonably by a NNEx utility function, it is natural to assume that, when he moves to $y_{ki} = x_{ki}$, his behaviour will be modelled by a NZEx utility function, so that his utility of wealth drops slightly. The process may be represented as $NNEx \rightarrow NZEx$ and is continuous. But switching from an NNEx to a Linear utility function will in almost all cases introduce a discontinuity in utility at the point where $y_{ki} = x_{ki}$.

The only case where such a discontinuity will be avoided is when

$$h_i = -(g_i w_i + 1) \quad (48)$$

Now the utility of wealth under the Linear utility function will match the utility under the NZEx, namely -1 , as may be seen by substituting Eq. (48) into Eq. (47), giving $h_i = -(g_i w_i + 1)$ for all w_i , an equation that will be seen to be identical to Eq. (24) that defines the NZEx utility function.

It can be shown by substituting Eq. (47) into Eq. (A.9) that the injury offset under the Linear utility function will be $a_i = g_i x_{ki}$. Since, $dU_i/dw_i = g_i$, it follows from Eq. (A.10) that $m_{ki} = x_{ki}$. Thus the Linear utility function gives the same marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , as the NZEx utility function. However it will offer the same seamless change in utility in the transfer from the region where the NNEx utility function is valid only when it is forced via the constraint of equation (48) to mimic exactly the behaviour of the NZEx utility function: it has turned into the NZEx utility function.

It may be concluded that the PPEX utility function is less valid than the NPEX utility function and that the Linear utility function has to be forced into the form of the NZEx utility function to attain the same validity as the NZEx utility function.

But are the preferred utility functions just discussed, NZEx and NPEX, themselves valid? Criticisms were made in Thomas

and Vaughan (2015b) of the two utility functions on the basis that they assume respondents who are either anti-materialist or at least indifferent to wealth. Moreover, the NZEx and the NPEX utility functions imply either risk-neutral or risk-seeking decision making, and it is difficult to believe that large numbers of the population would adopt anything other than a risk-averse stance when taking decisions on safety, specifically in the Carthy study on two serious injuries, involving hospitalisation and a recovery period of 3–4 months in the one case and fully 18 months in the other. Indeed such a stipulation on risk-aversion is included in Section 1 of Carthy et al. (1999), where it is required that the individual

“is financially risk-averse so that $(\forall w) U'(w) > 0, U''(w) < 0$ where $U''(w)$ denotes the second derivative of $U(\cdot)$ ”

Substituting the mathematical conditions just listed into Eq. (2) shows that the individual's risk-aversion, ε_i , is positive. The statement above by the Carthy authors is thus equivalent to the requirement that the individual's risk-aversion should be positive: $\varepsilon_i > 0$. (Both the PPEX and the Linear utility functions are, of course, inconsistent with this condition).

The Exponential utility functions, NPEX and NZEx, parallel the behaviour of the Constrained Power utility when $MAC < MAP$ (Thomas and Vaughan, 2015a, Section 5.4) in the following respect. In being forced to model risk-neutral ($\varepsilon_i = 0$) or risk-seeking behaviour ($\varepsilon_i < 0$), the Exponential utility function is pushed beyond its region of applicability to safety decisions. The conclusion here must be the same as for the Constrained Power utility function in such circumstances, namely that the results generated by the NPEX and NZEx utility functions should not be regarded as reliable.

Since the PPEX and Linear utility functions have less validity than the NPEX and NZEx utility functions in modelling the individual's safety decision making, it follows that results from the PPEX and Linear utility functions discussed by Jones-Lee and Loomes (2015) cannot be relied upon and should not be used.

7. Joining utility functions to model an individual's utility of wealth

The Exponential utility functions discussed in Sections 3 and 4 above are defined fully, so that it is possible to deduce the individual's utility of wealth for all possible wealths in the range: $w_i : 0 \leq w_i < \infty$ as soon as β_i is found, for example from Eq. (14) above. In particular, no arbitrary wealth threshold, $\beta^{(i)}$, is subtracted from the individual's wealth before applying the utility function. This was the stance adopted by Carthy et al. (1999) when they applied the Negative Exponential utility function. Consistency across the 4 utility functions used in the Carthy study (Constrained Power, Logarithmic, Negative Inverse and Negative Exponential) would seem to imply that any wealth threshold must be zero: $\beta^{(i)} = 0$, as explained in Thomas and Vaughan, 2015b. However Jones-Lee and Loomes (2015) have now made it clear that they fully intended to introduce a positive wealth threshold, $\beta^{(i)} > 0$, of indeterminate size into their specification of the Constrained Power, Logarithmic and Negative Inverse utility functions. (For reference, the formulae for these utility functions with a wealth threshold, $\beta^{(i)}$, included are given in Section 7.1, Eq. (49), Section 7.2, Eq. (55) and Section 7.3, Eq. (57) respectively.)

In fact, the results for VPF will be independent of $\beta^{(i)}$, and so it is not surprising that the same results were derived in Thomas and Vaughan (2015a) after setting $\beta^{(i)} = 0$ as found by

Carthy et al. (1999), who apparently assumed a positive but otherwise unspecified value for the wealth threshold, $\beta^{(i)}$. Paralleling the case with the two groups of Exponential utility functions discussed in Section 6, Negative and Positive, there is no way of deducing from the inputs, MAP, x_{ki} , and MAC, y_{ki} , and the output, m_{ki} , whether the utility function employed a wealth threshold or none.

Setting $\beta^{(i)} = 0$ would appear to be the most natural way of modelling human appreciation of wealth, since the utility function will then apply to any non-negative value of wealth, meaning that the individual will place some value on his wealth even if it is very low. However, the intention of Jones-Lee and Loomes (2015) is that $\beta^{(i)}$ should be regarded as a non-negative parameter subject only to the proviso that its size should allow the computation of the utility of wealth for a wealth of $w_{i0} - \beta^{(i)} - x_{ki}$. Thus for the Constrained Power utility function it is required that $w_{i0} - \beta^{(i)} - x_{ki} \geq 0$, leading to the closed limits on $\beta^{(i)}$: $0 \leq \beta^{(i)} \leq w_{i0} - x_{ki}$. The impossibility of computing the utility of zero wealth under the Logarithmic and Negative Inverse utility functions, on the other hand, means that for these utility functions the upper limit is open: $0 \leq \beta^{(i)} < w_{i0} - x_{ki}$.

The effect of any positive wealth threshold, $\beta^{(i)} > 0$ is to limit the application of the Constrained Power to wealths $w_i \geq \beta^{(i)}$ and leave the individual's utility undefined for $0 \leq w_i < \beta^{(i)}$. Similarly the Logarithmic and Negative Inverse utility functions will apply only to wealths $w_i > \beta^{(i)}$ and leave the individual's utility undefined for $0 \leq w_i \leq \beta^{(i)}$. But Jones-Lee and Loomes (2015) wish to counter the charge of lack of definition at and below $w_i = \beta^{(i)}$, and now suggest that a second utility function of a different type could be used to cover the region below the individual's wealth threshold, $\beta^{(i)}$. It will be shown that this is an unrealistic strategy for the Constrained Power utility function and an impossible strategy for the other two.

The case when the Constrained Power constitutes the main utility function under consideration, for convenience named here “the first utility function”, is explored in the next subsection.

7.1. The first utility function is the Constrained Power utility function

The properties of the Constrained Power utility are derived in detail in Appendix B. Allowing for a wealth threshold, $\beta^{(i)}$, and choosing $g_i = 1$ and $h_i = 0$, the first utility function has the form:

$$U_{1i}(w_i) = (w_i - \beta^{(i)})^{1-\varepsilon_i} \quad w_i \geq \beta^{(i)} \quad (49)$$

Now let us assume that the second utility function that is to be joined to it is Logarithmic:

$$U_{2i}(w_i) = g_i \ln w_i + h_i \quad 0 < w_i \leq \beta^{(i)}; \quad g_i > 0 \quad (50)$$

The retention of the parameters, g_i and h_i , conforms to the implication from Jones-Lee and Loomes that the flexibility of the positive linear transformation should be used to allow the second utility function, valid over $0 < w_i \leq \beta^{(i)}$, to match the first utility function, valid for $w_i \geq \beta^{(i)}$, at the cross-over point, $w_i = \beta^{(i)}$. Matching the two utilities at the cross-over wealth requires that:

$$h_i = -g_i \ln \beta^{(i)} \quad (51)$$

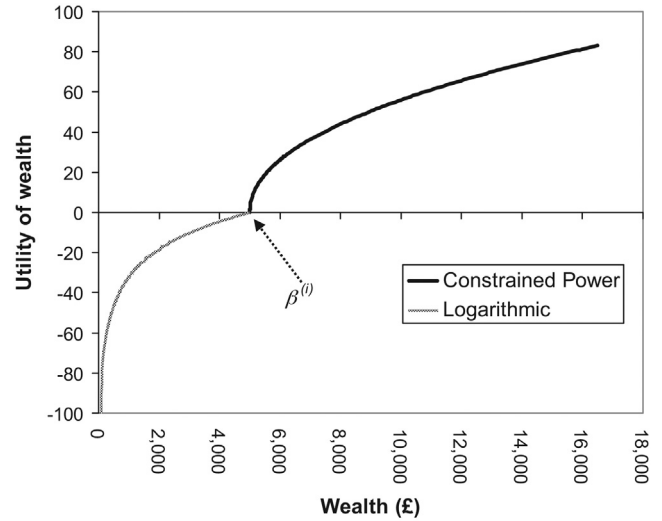


Fig. 5 – Joining a Logarithmic utility function to a Constrained Power utility function. $\beta^{(i)} = £5000$.

Take the example of Respondent 119 in the Carthy study, who set his MAP, x_{Xi} , to avert injury X as $x_{Xi} = £7,500$ and his MAC, y_{Xi} , for enduring injury X as $y_{Xi} = £25,000$ so that, from Eq. (B.10), $\varepsilon_i = 1 - \frac{\ln 2}{\ln(1+y_{Xi}/x_{Xi})} = 0.5273$. For the purposes of illustration assume that $\beta^{(i)} = £5,000$ (the precise value will not affect the argument). Selecting $g_i = 20$ for example (again the precise value will not affect the argument), Eq. (51) yields $h_i = -170.3$. Fig. 5 shows the resulting combined utility function, and the join is very obvious, with a sudden, enormous increase in the individual's utility of wealth at $w_i = \beta^{(i)}$. But what mechanism could be at work to generate such a huge increase in utility at some mid-range wealth, $\beta^{(i)}$?

In fact, in introducing the concept of utility, Daniel Bernoulli (1738) considered a similar problem:

“though a poor man generally obtains more utility than does a rich man from an equal gain, it is nevertheless conceivable, for example, that a rich prisoner who possesses two thousand ducats but needs two thousand ducats more to repurchase his freedom, will place a higher value on a gain of two thousand ducats than does another man who has less money than he.”

and we might comment that he would value a wealth of 4000 ducats hugely more than 3999. But we believe that Bernoulli's answer remains valid:

“Though innumerable examples of this kind may be constructed, they represent exceedingly rare exceptions. We shall, therefore, do better to consider what usually happens.”

Viewed in this light, the pronounced kink is obviously a distortion of the normal run of things. However, it is unavoidable under the proposition of Jones-Lee and Loomes because, although we can make $U_{1i}(w_i) = U_{2i}(w_i)$ at $w_i = \beta^{(i)}$, the first derivatives cannot match. For, in the case when the 1st utility function is of the Constrained Power type and the second is of Logarithmic form:

$$\frac{dU_{2i}}{dw_i} = \frac{g_i}{w_i} \rightarrow \frac{g_i}{\beta^{(i)}} \quad \text{as } w_i \rightarrow \beta^{(i)} \quad (52)$$

but

$$\frac{dU_{1i}}{dw_i} = \frac{1 - \varepsilon_i}{(w_i - \beta^{(i)})^{\varepsilon_i}} \rightarrow \infty \text{ as } w_i \rightarrow \beta^{(i)} \quad (53)$$

It is clearly impossible for the second utility function to match the infinite slope of the first utility function at $w_i = \beta^{(i)}$ whatever finite value of g_i is chosen. The same problem will occur if the second utility function is either the Negative Inverse, when $dU_{2i}/dw_i = g_i/(\beta^{(i)})^2$, or the NNEx: $dU_{2i}/dw_i = g_i\beta_i e^{-\beta_i\beta^{(i)}}$, or even another Constrained Power: $dU_{2i}/dw_i = g_i(1 - \varepsilon_i)/(\beta^{(i)})^{\varepsilon_i}$. None of these derivatives can match the infinite slope of the first utility function at $w_i = \beta^{(i)}$. The distorting kink will thus occur for any wealth threshold apart from $\beta^{(i)} = 0$.

The presence of the kink incorporating an infinite slope at the wealth threshold means that the utility of wealth below that threshold cannot be defined by any conventional utility function and so must remain undefined. Unless the wealth threshold is zero, or very close (at, for example, the “quantum of wealth”, a low value, estimated at between £1 and £5 in Thomas et al. (2010d)), the sudden increase in utility of wealth at some positive wealth calls into question the validity of any first utility function that allows for a wealth threshold, $\beta^{(i)}$, other than zero or negligibly small.

Thus, despite the limited degree of matching achieved through Eq. (51), the process of joining two utility functions as suggested by Jones-Lee and Loomes will always produce a glaring anomaly when the first utility function is a Constrained Power. It is clear that the combined utility function cannot then be regarded as a reasonable representation of human preferences.

The only valid value of $\beta^{(i)}$ emerges as that which can remove the distorting kink, namely $\beta^{(i)} = 0$. Hence in Eq. (B.6) of Appendix B, the wealth difference, $\Delta w_{i0} - \beta^{(i)}$, is transformed into the wealth: $\Delta w_{i0} = w_{i0}$. This allows the individual's starting wealth, w_{i0} , to be found in the manner described in Thomas and Vaughan (2015a), namely $w_{i0} = x_{ki}$, $k = W, X$ (Eq. (B.11) of Appendix B).

Another perspective is offered if the view is taken that the general shape of the combined utility function of Fig. 5 would be roughly correct provided the distorting kink were removed. It is in fact possible to fit the utility function

$$U_i(w_i) = 1.585w_i^{1-\varepsilon_i} - 78.04 \quad w_i \geq 0; \quad \varepsilon_i = 0.5273 \quad (54)$$

which takes out the kink and provides the reasonably good match shown in Fig. 6.

The utility function of Eq. (54) has a wealth threshold of zero, and so $\Delta w_{i0} = w_{i0} = x_{ki}$. Thus, for example, the wealth of Respondent 119 is £7500 under injury X, the figure found previously by applying the Constrained Power utility function as in Thomas and Vaughan (2015a).

7.2. The first utility function is the Logarithmic utility function

The first utility function now takes has the form:

$$U_{1i}(w_i) = \ln(w_i - \beta^{(i)}) \quad w_i > \beta^{(i)} \quad (55)$$

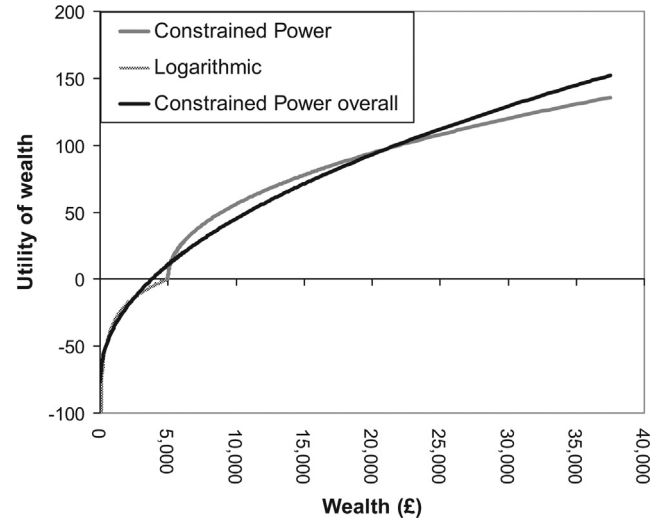


Fig. 6 – Fitting a single Constrained Power to the two utility functions.

However, any attempt to fit a contiguous second utility function, $U_{2i}(w_i)$, to cover the region $0 < w_i \leq \beta^{(i)}$, with $U_{2i}(\beta^{(i)}) = U_{1i}(\beta^{(i)})$, is doomed to failure, as

$$U_{1i}(w_i) = \ln(w_i - \beta^{(i)}) \rightarrow -\infty \text{ as } w_i \rightarrow \beta^{(i)} \quad (56)$$

It is impossible to find a second utility function of any finite form that will fit underneath the first utility function when the first utility function is Logarithmic. Fig. 7 illustrates this insuperable problem.

7.3. The first utility function is the Negative Inverse utility function

The first utility function now takes has the form:

$$U_{1i}(w_i) = -\frac{1}{w_i - \beta^{(i)}} \quad w_i > \beta^{(i)} \quad (57)$$

Once again any attempt to fit a contiguous second utility function, $U_{2i}(w_i)$, to cover the region $0 < w_i \leq \beta^{(i)}$, with $U_{2i}(\beta^{(i)}) = U_{1i}(\beta^{(i)})$, is doomed to failure, as

$$U_{1i}(w_i) = -\frac{1}{w_i - \beta^{(i)}} \rightarrow -\infty \text{ as } w_i \rightarrow \beta^{(i)} \quad (58)$$

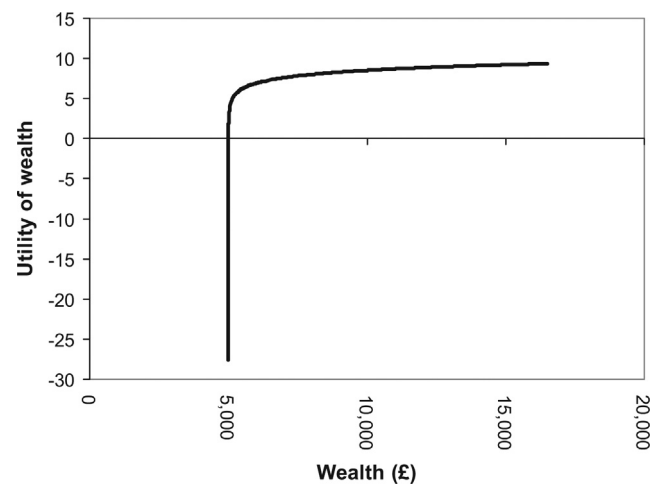


Fig. 7 – Logarithmic utility of wealth for $\beta^{(i)} = £ 5000$.

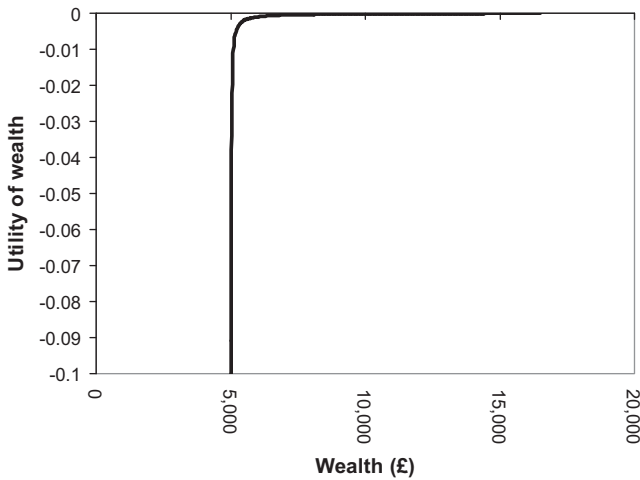


Fig. 8 – Negative Inverse utility of wealth for $\beta^{(i)} = £5000$.

As with the Logarithmic, it is impossible to find a second utility function of any finite form that will fit underneath the first utility function when the first utility function is the Inverse Negative. Fig. 8 illustrates the point.

7.4. General problems with joining utility functions

It is understandable that the first increments of wealth just above zero should be valued very highly, providing a secure basis for ensuring the fundamentals of life are met. However, it appears reasonable to assume that subsequent increments would be valued less. This situation, which occurs when the wealth threshold is zero, is fully explicable on the basis of economics and common sense.

By comparison, any generally applicable model whereby the utility the individual derives from his wealth should increase dramatically at some wealth, $\beta^{(i)}$, must be regarded as unrealistic. Even without the detailed mathematical confirmation presented above, it is clear that there is no realistic possibility, except perhaps in special circumstances, of joining two utility functions for wealth in the way suggested by Jones-Lee and Loomes (2015). The sharp and discontinuous rise in the utility of wealth calculated by the 1st utility function at the wealth threshold can be characteristic only of the very first increment of wealth not a mid-range wealth. We may conclude that a positive wealth threshold will lead to a utility that is undefined below the wealth threshold and unrealistic above it.

It is notable that when the first utility function is either Logarithmic or Negative Inverse (Sections 7.2 and 7.3), the individual with a positive wealth threshold, $\beta^{(i)}$, must be unmoved by any change in his wealth below this level, since his utility must at all times be less than or equal to minus infinity. Thus his assets may double in value but he will be indifferent to this good fortune if his resultant wealth stays below the threshold. Such a situation can be a fair model of reality only when the threshold is set at the “quantum of wealth”, which has been estimated to be of the order of £1.

Using a single utility function in the way shown in Fig. 6, for example, avoids the unrepresentative kink and offers the best chance of mirroring reality and coping with the full range of wealths that a person might need to evaluate. But, of course, this approach implies $\beta^{(i)} = 0$.

The arguments above make it clear that introducing an incompletely defined wealth threshold, $\beta^{(i)} > 0$, into the utility

function brings insuperable problems for the notion that the utility of the individual's wealth below $\beta^{(i)}$ can be defined by another utility function operating over $0 < w_i < \beta^{(i)}$.

As noted above, the value wealth threshold, $\beta^{(i)}$, makes no difference to the calculation of the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth, where m_{ki} is an intermediate parameter in the calculation of the VPF using the two-injury chained model. So the question arises, what purpose does it serve to introduce such an extra parameter? It is also relevant to question why such a parameter introduced into the model should be left undefined, or at least free to range over a potentially large interval. If it was important to the Carthy study, why were the respondents not asked the value they would put on it?

It might be thought that the action of specifying that some positive wealth threshold, $\beta^{(i)}$, should exist for each individual but at the same time leaving it incompletely specified, might prevent what Chilton et al. (2015) and Jones-Lee and Loomes (2015) describe as the “ridiculously low levels” of implied wealth for the respondents in the Carthy study becoming apparent. But this is not the case, as will now be shown.

8. The average wealth of the respondents when it is assumed that there is some positive wealth threshold

Although it has been shown untenable in the previous section, let us assume purely for the sake of argument and illustration that a strictly positive and non-negligible wealth threshold, $\beta^{(i)}$, can exist, as proposed by Jones-Lee and Loomes. The effect of a non-zero wealth threshold will be considered for each of the three utility functions used in Carthy et al. (1999) that embody a predictive capability for the individual's wealth, namely the Constrained Power, the Logarithmic and the Negative Inverse. (The Negative Exponential utility function, also used in the Carthy study, is exceptional in this regard, in that the individual's wealth cannot be inferred from his MAP and MAC values.)

8.1. Constrained Power utility function

When a wealth threshold is included, the Constrained Power utility function will take the form:

$$U_i(w_i) = \Delta w_i^{1-\varepsilon_i} = \Delta w_i^{s_i} \quad (59)$$

where Δw_i is the individual's wealth less his supposed wealth threshold:

$$\Delta w_i = w_i - \beta^{(i)} \quad 0 \leq \beta^{(i)} \leq w_{i0} - x_{ki} \quad (60)$$

with the constraint on $\beta^{(i)}$ appended to Eq. (60) reflecting the computability requirement of Eq. (59) for non-integer $s_i = 1 - \varepsilon_i$. The assumption inherent in the Constrained Power utility function is that the injury offset will take its maximum permissible value, which, as shown in Appendix B, will occur when

$$\Delta w_{i0} = x_{ki} \quad k = W, X \quad (B.6)$$

Since the wealth threshold $\beta^{(i)}$ lies in the interval: $0 \leq \beta^{(i)} \leq w_{i0} - x_{ki}$, it may be represented by:

$$\beta^{(i)} = \lambda_i (w_{i0} - x_{ki}) \quad 0 \leq \lambda_i \leq 1 \quad (61)$$

Combining Eqs. (60), (61) and (B.6) gives

$$w_{i0} - \lambda_i (w_{i0} - x_{ki}) = x_{ki} \quad (62)$$

or

$$(1 - \lambda_i) w_{i0} = (1 - \lambda_i) x_{ki} \quad (63)$$

so that

$$w_{i0} = x_{ki} \quad (64)$$

for any $\lambda_i : 0 \leq \lambda_i < 1$ and hence for any $\beta^{(i)} : 0 \leq \beta^{(i)} < w_{i0} - x_{ki}$. It may be seen that Eq. (64) yields the same result for the individual's wealth as given in Thomas and Vaughan (2015a), where it was assumed, on the grounds of consistency, that $\beta^{(i)} = 0$.

Thus the results for the implied wealth of the respondents given in Thomas and Vaughan (2015a) under the Constrained Power utility function remain the same. Hence the average wealths calculated for the respondents will also stay unchanged: £5252 under injury X and £1730 under injury W, compared with the UK average of £78,300 at the time of the survey (Matheson and Summerfield, 2000). As Jones-Lee and Loomes themselves admit, the implied wealths are “absurdly low” (Jones-Lee and Loomes, 2015).

There should not, of course, be one average value for injury W and a different average for injury X, with the two estimates differing by a factor of three. They should be the same. This failing constitutes another major problem for the Carthy study, sufficient on its own to invalidate the Carthy results.

The very low value of average respondent wealth embodied in these results from the two-injury chained method is unaffected by the introduction of a wealth threshold. These “absurdly low” values may be explained in one of 3 ways:

- (i) the selection of respondents failed to cover the full range of wealth as required, but was biased towards those with the lowest wealths, or
- (ii) the selection was fair but the method gives incorrect answers, or
- (iii) a mixture of (i) and (ii).

Whichever way, it is possible to conclude that the VPF arising from the two-injury chained method is unsubstantiated.

Jones-Lee and Loomes (2015) claim that the choice of respondents was “broadly representative” based on the collected “demographic data concerning income, age and social class”. From this we might conclude that it is the method is at fault. But we also learn that it will never be possible to compare the predicted and actual wealths for the Carthy study. This is because, while they gathered information on the respondents' incomes, presumably by asking them directly or getting others to do so, the Carthy authors made no attempt to find the individuals' wealths that are key to justifying any VPF figure:

“we certainly did not have the temerity to question respondents about their overall wealth”

We will therefore never know whether the method or the data were at fault, but the fact that other failings have been identified in the methodology suggests that the method itself represents a significant source of error.

8.2. Logarithmic utility function

If, as proposed by Jones-Lee and Loomes, a non-zero wealth threshold, $\beta^{(i)}$, can exist, then the Logarithmic utility function becomes:

$$U_i(w_i) = \ln \Delta w_i \quad (65)$$

Eq. (65) replaces Eq. (B.13) of Thomas and Vaughan (2015a), derived from Bernoulli (1738). Applying Eq. (65) to Eq. (A.8) then produces

$$\ln(\Delta w_{i0} - x_{ki}) + \ln(\Delta w_{i0} + y_{ki}) = 2 \ln(\Delta w_{i0}) \quad (66)$$

which implies

$$\Delta w_{i0} = \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} \quad (67)$$

where

$$\Delta w_{i0} = w_{i0} - \beta^{(i)} \quad 0 \leq \beta^{(i)} < w_{i0} - x_{ki} \quad (68)$$

The constraint on wealth threshold, $\beta^{(i)}$, reflects the computability requirements of Eq. (65). Since it lies in the semi-closed range: $0 \leq \beta^{(i)} < w_{i0} - x_{ki}$, $\beta^{(i)}$ may be represented by:

$$\beta^{(i)} = \lambda_i (w_{i0} - x_{ki}) \quad 0 \leq \lambda_i < 1 \quad (69)$$

Combining Eqs. (67), (68) and (69) gives

$$w_{i0} - \lambda_i (w_{i0} - x_{ki}) = \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} \quad (70)$$

which may be re-arranged to

$$w_{i0} (1 - \lambda_i) = \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} - \lambda_i x_{ki} = (1 - \lambda_i) \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} + \lambda_i \frac{x_{ki}^2}{y_{ki} - x_{ki}} \quad (71)$$

Hence

$$\begin{aligned} w_{i0} &= \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} + \frac{\lambda_i}{1 - \lambda_i} \frac{x_{ki}^2}{y_{ki} - x_{ki}} \\ &= \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} + m_i \frac{x_{ki}^2}{y_{ki} - x_{ki}} \end{aligned} \quad (72)$$

where m_i is the ratio of λ_i to its complement, $1 - \lambda_i$:

$$m_i = \frac{\lambda_i}{1 - \lambda_i} \quad (73)$$

The proposition of Jones-Lee and Loomes is that the wealth threshold, $\beta^{(i)}$, lies within a range but is otherwise unknown. This makes it reasonable to model it as a random variable distributed over its allowable range: $0 \leq \beta^{(i)} < w_{i0} - x_{ki}$, which may be achieved by taking λ_i to be random over the range 0 to 1: $0 \leq \lambda_i < 1$. This requires that m_i and w_{i0} should be replaced by the random variables, M_i and W_{i0} , so that:

$$W_{i0} = \frac{x_{ki} y_{ki}}{y_{ki} - x_{ki}} + M_i \frac{x_{ki}^2}{y_{ki} - x_{ki}} \quad (74)$$

where,

$$M_i = m_i (A_i) \quad (75)$$

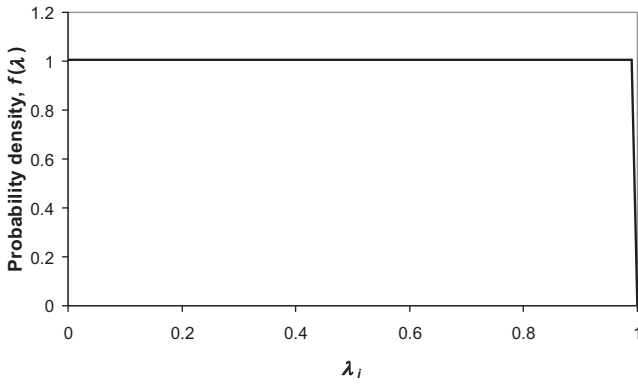


Fig. 9 – Probability density, $f(\lambda_i)$.

Importantly, the expected starting value of the wealth, $E(W_{i0})$, of respondent i may be found by applying the expectation operator to Eq. (74):

$$E(W_{i0}) = \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + \frac{x_{ki}^2}{y_{ki} - x_{ki}} E(M_i) \quad (76)$$

With nothing more than its range to go on, a first thought might be to model λ_i as uniform over $0 \leq \lambda_i < 1$, but the term, $1 - \lambda_i$, in the denominator of Eq. (73) would mean that $m_i \rightarrow \infty$ as $\lambda_i \rightarrow 1$, making $E(M_i)$ infinite also. The impossibility of infinite wealth rules out the uniform distribution as a model for the behaviour of λ_i , but a quasi-uniform distribution may be retained, whereby the probability density, $f(\lambda_i)$, is uniform up to some point, $\lambda_i = \lambda_a$, with the probability density then decreasing linearly to zero as $\lambda_i \rightarrow 1$:

$$f(\lambda_i) = \begin{cases} \frac{2}{1 + \lambda_a} & \text{for } 0 \leq \lambda_i \leq \lambda_a \\ \frac{2}{1 - \lambda_a^2} (1 - \lambda_i) & \text{for } \lambda_a \leq \lambda_i < 1.0 \end{cases} \quad (77)$$

Setting $\lambda_a = 0.99$ means that the probability distribution for the wealth threshold, $\beta^{(i)}$, is uniform as far as 99% of its maximum possible value, after which it reduces linearly to zero over the final 1% of values. See Fig. 9.

The procedure for finding $E(M_i)$ is given in Appendix C and at $\lambda_a = 0.99$ produces a value of $E(M_i) = 4.63$. This allows the expected wealth, $E(W_{i0})$, to be found for each respondent, using Eq. (76). The average, \bar{W}_0 , of the expected starting wealths, $E(W_{i0})$, may then be found by summing the expected wealths of all respondents and dividing by their number. The expected average wealth, $E(\bar{W}_0)$, amongst all the N_s respondents in the sample is found by applying the expectation operator, which will yield the same value:

$$E(\bar{W}_0) = \frac{1}{N_s} \sum_{i=1}^{N_s} E(W_{i0}) \quad (78)$$

8.3. Negative Inverse utility function

Under the proposition of Jones-Lee and Loomes that a non-zero wealth threshold, $\beta^{(i)}$, can exist, the new equation defining the Negative Inverse utility function is

$$U_i(w_i) = -\frac{1}{\Delta w_i} \quad (79)$$

where, $\Delta w_i = w_i - \beta^{(i)}$. Applying Eq. (79) to Eq. (A.8) then produces

$$-\frac{1}{\Delta w_{i0} - x_{ki}} - \frac{1}{\Delta w_{i0} + y_{ki}} = -\frac{2}{\Delta w_{i0}} \quad (80)$$

which implies

$$\Delta w_{i0} = \frac{2x_{ki}y_{ki}}{y_{ki} - x_{ki}} \quad (81)$$

Representing the wealth threshold $\beta^{(i)}$ by Eq. (69) as before, combining Eqs. (69), (80) and (81) gives

$$w_{i0} - \lambda_i(w_{i0} - x_{ki}) = \frac{2x_{ki}y_{ki}}{y_{ki} - x_{ki}} \quad (82)$$

which may be re-arranged to

$$w_{i0}(1 - \lambda_i) = \frac{2x_{ki}y_{ki}}{y_{ki} - x_{ki}} - \lambda_i x_{ki} = (2 - \lambda_i) \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + \lambda_i \frac{x_{ki}^2}{y_{ki} - x_{ki}} \quad (83)$$

Hence

$$w_{i0} = \frac{2 - \lambda_i}{1 - \lambda_i} \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + \frac{\lambda_i}{1 - \lambda_i} \frac{x_{ki}^2}{y_{ki} - x_{ki}} \quad (84)$$

But

$$\frac{2 - \lambda_i}{1 - \lambda_i} = \frac{2 - \lambda_i - \lambda_i + \lambda_i}{1 - \lambda_i} = 2 \frac{1 - \lambda_i}{1 - \lambda_i} + \frac{\lambda_i}{1 - \lambda_i} = 2 + m_i \quad (85)$$

where Eq. (73) has been used in the last step. Hence

$$\begin{aligned} w_{i0} &= (2 + m_i) \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + m_i \frac{x_{ki}^2}{y_{ki} - x_{ki}} \\ &= 2 \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + m_i \left(\frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + \frac{x_{ki}^2}{y_{ki} - x_{ki}} \right) \end{aligned} \quad (86)$$

Proceeding as before, assume that λ_i is random, so that:

$$W_{i0} = 2 \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + M_i \left(\frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + \frac{x_{ki}^2}{y_{ki} - x_{ki}} \right) \quad (87)$$

while

$$E(W_{i0}) = 2 \frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + E(M_i) \left(\frac{x_{ki}y_{ki}}{y_{ki} - x_{ki}} + \frac{x_{ki}^2}{y_{ki} - x_{ki}} \right) \quad (88)$$

Using Eq. (88) in Eq. (78) then gives the expected average wealth under the Negative Inverse utility function.

8.4. Results on apparent wealth of respondents

The results for the three utility functions are given in Table 2, using data listed in the document “Individual Responses & VOSLS-PEG Study”, HSE/Peg/Nov 1997/SC, with zero and negative wealths excluded. The average wealths are found for the Constrained Power utility function and are, of course, the low values reported previously in Thomas and Vaughan (2015a). Meanwhile the random variable model applied to both the Logarithmic and the Negative Inverse utility functions produces the expected average wealths listed. (The two values per utility function represent the different values derived

Table 2 – Respondents' average wealth and expected average wealth under the three utility functions: Constrained Power, Logarithmic and Negative Inverse.

	Average wealth (£)	Expected average wealth (£)	
Utility function	Constrained Power	Logarithmic	Negative Inverse
Injury X	5252	29,707	77,902
Injury W	1730	4811	14,703

depending on whether injury W is considered or injury X. As noted previously, this represents a clear and obvious flaw in the two-injury chained method.)

While now one of the results (Negative Inverse, injury X) comes close to the average wealth, £78,300, of a UK adult in 1997 (Matheson and Summerfield, 2000), 5 out of the 6 figures for average wealth/expected average wealth lie well below that figure. Indeed, no reliance can be placed on the figure produced by the Negative Inverse utility function under injury X as, under injury W, it produces a sum that is 5 times less, when the two values should be the same.

This variability continues as a further major and invalidating problem for any argument in favour of introducing a wealth threshold. Common sense suggests that the predictions for average wealth made by different methods for the same sample population should come out the same or at least close to each other. But, as illustrated in the linear graph of Fig. 10, the predictions are very widely spread, between close to zero assets and the average wealth of the UK adult. Modifying the statistical model used to find the expected average values for the Logarithmic and Negative Inverse utility functions by allowing the uniform distribution to cover 99.9% or 99.999% of the possible range rather than 99% would merely accentuate the spread, partly (but not solely) because the values coming from the Constrained Power utility function are firmly anchored at £1730 and £5250.

So even if all the contrary evidence of Section 7 were to be ignored and the proposition of an incompletely defined wealth

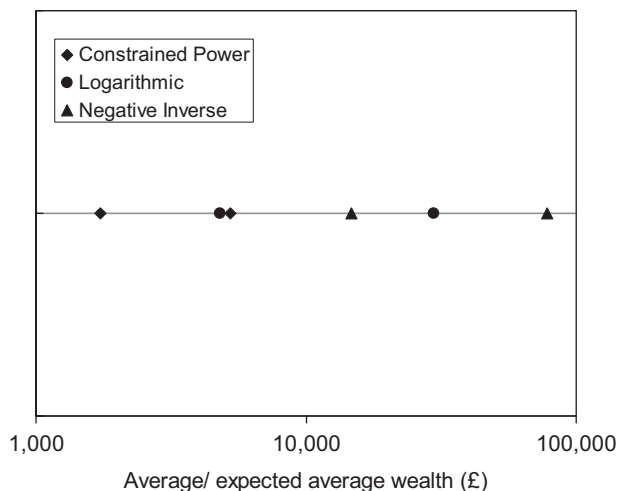


Fig. 10 – Line graph showing average wealth or expected average wealth under the three utility functions: Constrained Power, Logarithmic and Negative Inverse. (The two points per utility function represent the different values derived depending on whether injury W is considered or injury X, indicating an obviously flaw in the two-injury chained method).

threshold were to be accepted, the gross variability of the figures for predicted average wealth would be sufficient in itself to show that the results produced by the two-injury chained method had been falsified in the sense described by Sir Karl Popper (Popper, 1934).

The results of the Carthy study are, of course, additionally invalidated by the low values of average wealth relative to the actual UK average at the time, as derived under the three utility functions: Constrained Power, Logarithmic and Negative Inverse. It is notable that the very low levels of wealth found under the Constrained Power utility function are unchanged by the addition of a wealth threshold.

9. Discussion on the use of utility functions

9.1. The duality of Exponential utility functions

As shown at the end of Section 4.1, the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , will always be the same for the Positive Exponential utility function as for the Negative Exponential utility function for any combination of MAP and MAC, (x_{ki}, y_{ki}) . Combining this fact with the relationships between the ranges of the parameter, $c_{ki} = y_{ki}/x_{ki}$, and the different categories of Exponential utility functions shown in Table 1, it is clear that there are 3 pairs of Exponential utility function that produce the same answer for the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth, w_i , given the same MAP, x_{ki} , and MAC, y_{ki} . Thus it is impossible to distinguish between the result of applying an NPEx utility function to find m_{ki} from x_{ki} and y_{ki} when $x_{ki} > y_{ki}$, as used by Thomas and Vaughan (2015a), and the result of applying a PPEx utility function, as used by Jones-Lee and Loomes (2015). In addition, the NZEx, the PZEx and the Linear utility functions will give the same value of m_{ki} from x_{ki} and y_{ki} at the point where $x_{ki} = y_{ki}$, so that it is again impossible to decide, based on the evidence of m_{ki} , x_{ki} and y_{ki} , which utility function has been used.

The duality between the Negative Exponential and Positive Exponential utility functions when they are used in the calculation of the marginal rate of substitution, m_{ki} , of wealth for non-injury probability was, in fact, foreshadowed, at least in part, by the Carthy authors, as explained in Section 5 above. On the other hand, it is possible that Jones-Lee and Loomes (2015) were not aware of the full degree of duality involved. Appendix C explains its significant extent.

For the avoidance of doubt, the same value for the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth will be produced whether cases where $c_{ki} \leq 1$ are calculated using the PPEx and Linear utility functions or the NPEx and NZEx utility functions. The choice makes no difference to the ultimate answer.

However, if the NNEx, the canonical form of the Negative Exponential utility function, is chosen as the default utility function, continuity of utility requires that the NZEx utility function be chosen for the case when MAP=MAC, and $c_{ki} = y_{ki}/x_{ki} = 1$, rather than either the PZEx or Linear utility functions. Similar grounds of continuity of utility apply for the case when MAP > MAC, $x_{ki} > y_{ki}$, requiring that NPEx should be chosen rather than the PPEx utility function discussed by Jones-Lee and Loomes (2015).

But all the utility functions, NPEx, PPEx, NZEx, PZEx and Linear, applicable when MAC is less than or equal to MAP, $c_{ki} \leq 1$, indicate behaviour that is either risk-neutral

or risk-seeking. This suggests that none will be a realistic representation of a person's decision making when he is considering how much to spend on averting a fatality. A breakdown in the two-injury chained method is indicated, affecting about 30% of the respondents in the Carthy study (under injury X or injury W or both), as explained previously in [Thomas and Vaughan \(2015a\)](#).

9.2. The proposition that two utility functions may be joined to cover wealths less than the wealth threshold, $\beta^{(i)}$

It is impossible, as shown above, to fit a utility function below either the Logarithmic or Negative Inverse utility functions to cover wealths less than the wealth threshold, $\beta^{(i)}$. Moreover, while it is possible to join a contiguous utility function to a Constrained Power utility function with wealth threshold, the exercise is rendered unworkable by the pronounced kink that the procedure introduces when wealth is equal to the wealth threshold. An infinite slope in the utility function now occurs, which is a distortion of reality. Aside from special circumstances (e.g. Bernoulli's wealthy prisoner), a sudden and dramatic rise in the person's utility occurring at this point can be characteristic only of the first increments of wealth just above zero, when the individual's wealth is providing a secure basis for ensuring the fundamentals of life are met. It goes against economics and common sense that much later increments will acquire a dramatically higher value than those preceding them.

Thus, despite the suggestion to the contrary of Jones-Lee and Loomes, it is infeasible to provide coverage below the wealth threshold by fitting a second, auxiliary utility function below a first utility function that embodies a wealth threshold when the first utility function is one of: Constrained Power, Logarithmic and Negative Inverse. Setting a wealth threshold other than zero would result in utility being undefined below the wealth threshold and unrealistic above. Thus the application of two utility functions to model utility above and below a positive wealth threshold is ruled out unless the wealth threshold is either zero or negligibly low. The use of a single utility function emerges as the only realistic way of coping with the full range of wealths that a person might need to evaluate.

It needs to be borne in mind that, although Jones-Lee and Loomes have clarified that they assumed the existence of an incompletely specified positive wealth threshold, its presence or absence makes no difference to the calculated VPF results.

Including no wealth threshold in the utility function is, of course, a fully respectable and, indeed, the conventional approach to calculating utility, employed since the time of Daniel Bernoulli (1738) and continuing to the present day (U.K. Treasury, 2011). So, irrespective of whether or not Jones-Lee and Loomes employed a positive wealth threshold, there is no reason why the MAP and MAC data of the Carthy study should not be interpreted using conventional utility functions with a wealth threshold of zero. Not only does this entirely reputable and usual approach allow all the Carthy study's results on VPF to be reproduced, it also allows the wealth of each respondent to be deduced for each of the Constrained Power, the Logarithmic and the Negative Inverse utility functions. The very low average wealths reported in [Thomas and Vaughan \(2015a\)](#) that result stand as deductions that are entirely legitimate on this basis, irrespective of the arguments for a positive wealth threshold put forward by Jones-Lee and Loomes, arguments that have been shown in any case to be unsustainable.

9.3. Implications for calculating respondents' wealth when an incompletely specified wealth threshold is included

But what happens to the calculations of respondents' wealth if, despite the arguments against it, an incompletely specified positive wealth threshold is nevertheless included in the utility function? The analysis of Section 8 shows that the results given in [Thomas and Vaughan \(2015a\)](#) are entirely unchanged for the Constrained Power utility function. The implied levels of wealth described by [Jones-Lee and Loomes \(2015\)](#) as "ridiculously low" and "absurdly low" remain just as they were before. The average wealth of the respondents is still £5252 under injury X and £1730 under injury W, figures that should be compared with the UK average of £78,300 at the time.

A deterministic calculation is not possible in the case of the Logarithmic and Negative Inverse utility functions, but a random variable analysis allows the expected average wealth of the respondents to be estimated for the Carthy survey. The model uses a quasi-uniform probability distribution giving equal weight to 99% of the possible range of the wealth threshold, with a linearly decreasing probability distribution for the final 1%. This allows plausible estimates to be produced for the expected average wealth of the respondents for both the Logarithmic and Negative Inverse utility functions. The spread between the wealth under injury W and injury X widens markedly, but the very low values remain. While one value approaches the national average, the other 3 figures are 38%, 19% and 6% of that figure. Of course, all the values for average or expected average wealth should come out the same, but, as demonstrated in the line graph of [Fig. 10](#), they are widely separated.

The conclusion remains that one of the following three options must apply:

- the Carthy study failed to select respondents representing the full range of wealths in the UK
- the selection was fair but the method employed gives incorrect answers
- the anomalous results arose from a combination of both the factors listed above

The VPF produced is unsubstantiated whichever is the correct explanation.

10. Conclusions

The VPF figure in use for the past 16 years in the UK rests squarely on the value produced by the Carthy study. The current authors have shown several deficiencies in the method used by Carthy et al. in an earlier paper. The resulting exchange of comments between ourselves and various members of the team who worked on the Carthy study has not led to satisfactory answers being provided to the concerns we raised. The latest commentary from Jones-Lee and Loomes is a spirited second defence of the Carthy study, but the points they raise may be seen to be technical rather than fundamental, and they have, in any case, been rebutted in full. Hence this second defence should not be used as an excuse for continuing to use a VPF figure that has been shown to be based on a study that does not stand up to scrutiny.

We are pleased to have this opportunity to elucidate some interesting features of utility functions and, indeed, some

potential pitfalls in their application to the valuation of human life. We are, however, saddened that Loomes and Jones-Lee seem in their second defence to adopt an *ad hominem* approach in responding to our arguments and that they declare themselves unwilling to debate the issues further. We have no desire to take a similar line and are content to allow the logic of our case to speak for itself.

The salient fact is that the VPF derived from the two-injury chained method has been shown to be unsubstantiated and it is wholly unsatisfactory that it should be used as a benchmark for safety investment in the UK. We stand by our previous conclusion (Thomas and Vaughan, 2015a) that there is a need for active consideration of methods of valuing human life in the UK that offer an alternative to stated preference techniques. Ensuring that UK workers and public receive adequate protection from industrial and transport hazards means that a re-appraisal is needed urgently of other statistical methodologies that can provide guidance to owners, operators and regulators on decisions on safety.

Acknowledgement

This work was carried out in response to correspondence associated with Thomas and Vaughan (2015a), which reported on work some of which was carried out as part of the NREFS project, Management of Nuclear Risk Issues: Environmental, Financial and Safety, led by City University London and carried out in collaboration with Manchester, Warwick and The Open Universities and with the support of the Atomic Energy Commission of India as part of the UK–India Civil Nuclear Power Collaboration. The authors acknowledge gratefully the support of the Engineering and Physical Sciences Research Council (EPSRC) under grant reference number EP/K007580/1. The views expressed in the paper are those of the authors and not necessarily those of the NREFS project.

Appendix A. Relating the individual's starting utility of wealth to his utilities after paying to avert the injury and after receiving compensation for the injury suffered

As part of the two-injury chained method, individuals, i , are invited to consider two injuries X and W , and asked to estimate the maximum acceptable price (MAP, £) they would pay to avert the injury and the minimum acceptable compensation (MAC, £) they would take as compensation for enduring the injury. Let w_i be the wealth of an individual, i , and $U_i(w_i)$ be his utility of wealth when he is in good health. Now consider an injury, k , that will reduce the individual's utility of wealth to $I_{ki}(w_i)$, $k = W, X$. A simple model for the utility in the injured condition is suggested in the Carthy study:

$$I_{ki}(w_i) = U_i(w_i) - a_{ki} \quad a_{ki} \geq 0, \quad k = W, X \quad (A.1)$$

where, a_{ki} may be termed the injury offset, particular to the injury and to the individual.

The MAP of individual i associated with injury k , x_{ki} , will be reached when its payment will reduce the utility of the healthy individual to the level he would experience without paying it and suffering the injury in consequence. His wealth, w_i , will have fallen from its starting level, w_{i0} , to

$$w_i = w_{i0} - x_{ki} \quad k = W, X \quad (A.2)$$

and thus his utility will be $U_i(w_i) = U_i(w_{i0} - x_{ki})$. The MAP will correspond to the indifference point where his utility in good health but having paid out x_{ki} equals his utility under injury but with his original wealth:

$$U_i(w_{i0} - x_{ki}) = I_{ki}(w_{i0}) \quad k = W, X \quad (A.3)$$

In an analogous way, the MAC, y_{ki} , will be reached when the utility of the injured person has risen by virtue of the increase in wealth to the level it would have been in the absence of both injury and compensation. His new wealth will be

$$w = w_{i0} + y_{ki} \quad k = W, X \quad (A.4)$$

and his new utility, because he is injured, will be $I_{ki}(w_i) = I_{ki}(w_{i0} + y_{ki})$. The MAC will occur at the indifference point where

$$U_i(w_{i0}) = I_{ki}(w_{i0} + y_{ki}) \quad k = W, X \quad (A.5)$$

Using Eq. (A.1) in Eqs. (A.3) and (A.5), we achieve the equation pair applicable to individual, i , considering either injury W or injury X :

$$U_i(w_{i0} - x_{ki}) = U_i(w_{i0}) - a_{ki} \quad k = W, X \quad (A.6)$$

$$U_i(w_{i0} + y_{ki}) - a_{ki} = U_i(w_{i0}) \quad k = W, X \quad (A.7)$$

from which a_{ki} may be eliminated to give:

$$U_i(w_{i0} - x_{ki}) + U_i(w_{i0} + y_{ki}) = 2U_i(w_{i0}) \quad k = W, X \quad (A.8)$$

The injury offset may then be found as:

$$a_{ki} = U_i(w_{i0}) - U_i(w_{i0} - x_{ki}) \quad k = W, X \quad (A.9)$$

Under the model used in the Carthy study, the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth, w_i , may then be found from (see Thomas and Vaughan, 2015a, Appendix A):

$$m_{ki} = \frac{a_{ki}}{\left. \frac{dU_i}{dw_i} \right|_{w_i=w_{i0}}} \quad k = W, X \quad (A.10)$$

Appendix B. Application to the Constrained Power utility function with wealth threshold and subject to a positive linear transformation

Allowing for a wealth threshold, $\beta^{(i)}$, and a positive linear transformation gives the utility of wealth, w_i , for individual, i , under the Constrained Power utility function:

$$U_i(w_i) = g_i(w_i - \beta^{(i)})^{s_i} + h_i \quad (B.1)$$

where, the exponent is the complement of the risk-aversion, $s_i = 1 - \varepsilon_i$, and $g_i > 0$. Substituting into Eq. (A.8) gives:

$$g_i(w_{i0} - \beta^{(i)} - x_{ki})^{s_i} + h_i + g_i(w_{i0} - \beta^{(i)} + y_{ki})^{s_i} + h_i = 2g_i(w_{i0} - \beta^{(i)})^{s_i} + 2h_i \quad k = W, X \quad (B.2)$$

Cancelling g_i and h_i and putting

$$\Delta w_{i0} = w_{i0} - \beta^{(i)} \quad (\text{B.3})$$

gives:

$$(\Delta w_{i0} - x_{ki})^{s_i} + (\Delta w_{i0} + y_{ki})^{s_i} = 2\Delta w_{i0}^{s_i} \quad k = W, X \quad (\text{B.4})$$

Meanwhile substituting from Eq. (B.1) into Eq. (A.9) gives the injury offset:

$$\begin{aligned} a_{ki} &= g_i (w_{i0} - \beta^{(i)})^{s_i} + h_i - g_i (w_{i0} - \beta^{(i)} - x_{ki})^{s_i} - h_i \quad k = W, X \\ &= g_i (\Delta w_{i0}^{s_i} - (\Delta w_{i0} - x_{ki})^{s_i}) \end{aligned} \quad (\text{B.5})$$

The assumption behind the Constrained Power utility function is that the injury offset will take its maximum permissible value, which, by inspection of Eq. (B.5), will occur when

$$\Delta w_{i0} = x_{ki} \quad k = W, X \quad (\text{B.6})$$

Differentiation of Eq. (B.1) with respect to w_i gives:

$$\frac{dU_i}{dw_i} = g_i s_i (w_i - \beta^{(i)})^{s_i-1} = g_i s_i \Delta w_{i0}^{s_i-1} \quad (\text{B.7})$$

Substituting from Eqs. (B.5), (B.6) and (B.7) into Eq. (A.10) then gives the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth, w_i , at $w_i = w_{i0}$ and hence $\Delta w_i = \Delta w_{i0}$

$$m_{ki} = \frac{\Delta w_{i0}^{s_i} - (\Delta w_{i0} - x_{ki})^{s_i}}{s_i \Delta w_{i0}^{s_i-1}} = \frac{x_{ki}}{s_i} \quad k = W, X \quad (\text{B.8})$$

Moreover substituting from Eq. (B.6) into Eq. (B.4) gives:

$$(x_{ki} + y_{ki})^{s_i} = 2x_{ki}^{s_i} \quad k = W, X \quad (\text{B.9})$$

Taking logs produces the following expression for the exponent:

$$s_i = \frac{\ln 2}{\ln(1 + y_{ki}/x_{ki})} = 1 - \varepsilon_i \quad k = W, X \quad (\text{B.10})$$

Note that Eqs. (B.6), (B.8) and (B.10) will hold whatever permissible values of g_i , h_i and $\beta^{(i)}$ are used. An extra important result deriving from the definition, $\Delta w_{i0} = w_{i0} - \beta^{(i)}$, is that, if $\beta^{(i)} = 0$, the individual's starting wealth is determined from Eq. (5):

$$w_{i0} = x_{ki} \quad k = W, X \quad (\text{B.11})$$

Appendix C. Further equivalences of the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth

It has already been shown in Section 4.1 that the marginal rate of substitution of non-injury probability in place of wealth, m_{ki} , will be the same for the Positive Exponential utility function as for the Negative Exponential utility function for any combination of MAP and MAC, (x_{ki}, y_{ki}) . This Appendix will show further equivalent expressions for the parameter, m_{ki} .

C.1. Alternative expression for m_{ki} for the Negative Exponential utility function

Putting

$$d_{ki} = \frac{x_{ki}}{y_{ki}} \quad (\text{C.1})$$

allows Eq. (8) to be expanded in the alternative way to Eq. (10):

$$e^{\beta_i y_{ki} d_{ki}} + e^{-\beta_i y_{ki}} = e^{-\beta_i y_{ki}} + (e^{\beta_i y_{ki}})^{d_{ki}} = 2 \quad (\text{C.2})$$

Putting

$$\theta_{ki} = e^{-\beta_i y_{ki}} \quad (\text{C.3})$$

and so

$$\beta_i = -\frac{\ln \theta_{ki}}{y_{ki}} \quad (\text{C.4})$$

allows Eq. (C.2) to be rewritten:

$$\theta_{ki} - 2 + \frac{1}{\theta_{ki}^{d_{ki}}} = 0 \quad (\text{C.5})$$

Comparing Eq. (C.5) with Eq. (12), it is clear that θ_{ki} will be the same function of d_{ki} as ϕ_{ki} is of c_{ki} . Hence

$$\theta_{ki} = F(d_{ki}) = F\left(\frac{x_{ki}}{y_{ki}}\right) \quad (\text{C.6})$$

An alternative formulation for injury offset, a_{ki} , to that given in Eq. (A.9) may be found from Eq. (A.7):

$$a_{ki} = U_i(w_{i0} + y_{ki}) - U_i(w_{i0}) \quad (\text{C.7})$$

Substituting from the definition of a Negative Exponential utility function of Eq. (3) then gives:

$$\begin{aligned} a_{ki} &= -e^{-\beta_i(w_{i0}+y_{ki})} + e^{-\beta_i w_{i0}} = U_i(w_{i0})(e^{-\beta_i y_{ki}} - 1) \\ &= U_i(w_{i0})(\theta_{ki} - 1) \end{aligned} \quad (\text{C.8})$$

Meanwhile differentiating Eq. (3) with respect to wealth gives:

$$\frac{dU_i}{dw_i} = \beta_i e^{-\beta_i w_i} = -\beta_i U_i = \frac{\ln \theta_{ki}}{y_{ki}} U_i \quad (\text{C.9})$$

where, Eq. (C.4) has been used in the last step. Substituting from Eqs. (C.7) and (C.8) into Eq. (A.10) then gives the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth as:

$$m_{ki} = \frac{\theta_{ki} - 1}{\ln \theta_{ki}} y_{ki} \quad (\text{C.10})$$

which is an alternative but equivalent form to Eq. (18). That the two are equivalent follows from expanding Eq. (8) as:

$$e^{\beta_i x_{ki}} - 1 = -(e^{-\beta_i y_{ki}} - 1) \quad (\text{C.11})$$

so that, using Eqs. (11) and (C.3)

$$\phi_{ki} - 1 = -(\theta_{ki} - 1) \quad (\text{C.12})$$

Dividing both sides of this equation by β_i gives

$$\frac{\phi_{ki} - 1}{\beta_i} = -\frac{\theta_{ki} - 1}{\beta_i} \quad (C.13)$$

Now substituting for β_i from Eq. (14) into the left-hand side and from Eq. (C.4) into the right-hand side of Eq. (C.12) gives the desired result:

$$\frac{\phi_{ki} - 1}{\ln \phi_{ki}} x_{ki} = \frac{\theta_{ki} - 1}{\ln \theta_{ki}} y_{ki} \quad (C.14)$$

The expressions for m_{ki} in Eqs. (18) and (C.10) will thus give the same result.

Eq. (C.10) may be rewritten using Eq. (C.6) as

$$m_{ki} = \frac{F\left(\frac{x_{ki}}{y_{ki}}\right) - 1}{\ln F\left(\frac{x_{ki}}{y_{ki}}\right)} y_{ki} \quad k = W, X \quad (C.15)$$

and compared with the expression previously derived as Eq. (19), repeated below

$$m_{ki} = \frac{F\left(\frac{y_{ki}}{x_{ki}}\right) - 1}{\ln F\left(\frac{y_{ki}}{x_{ki}}\right)} x_{ki} \quad k = W, X \quad (19)$$

The two expressions will give the same result for every pair of MAP and MAC, (x_{ki}, y_{ki}) . It may be observed that the right-hand side of Eq. (19) may be converted into Eq. (C.15) by replacing y_{ki} by x_{ki} and x_{ki} by y_{ki} .

C.2. Alternative expression for m_{ki} for the Positive Exponential utility function

Using Eq. (C.1) allows Eq. (28) to be expanded in an alternative way to Eq. (29):

$$e^{-\beta_i d_{ki} y_{ki}} + e^{\beta_i y_{ki}} = e^{\beta_i y_{ki}} + (e^{-\beta_i y_{ki}})^{d_{ki}} = 2 \quad (C.16)$$

Put

$$x_{ki} = e^{\beta_i y_{ki}} \quad (C.17)$$

and so

$$\beta_i = \frac{\ln x_{ki}}{y_{ki}} \quad (C.18)$$

Substituting from Eq. (C.17) into Eq. (C.16) then gives

$$x_{ki} - 2 + \frac{1}{x_{ki}^{d_{ki}}} = 0 \quad (C.19)$$

Eq. (C.19) is of the same form in x_{ki} as Eq. (C.5) in θ_{ki} , and so we may write

$$x_{ki} = F(d_{ki}) = F\left(\frac{x_{ki}}{y_{ki}}\right) \quad (C.20)$$

Substituting from the definition of a Positive Exponential utility function of Eq. (5) into Eq. (C.7) for injury offset gives:

$$a_{ki} = e^{\beta_i(w_{i0} + y_{ki})} - e^{\beta_i w_{i0}} = U_i(w_{i0}) (e^{\beta_i y_{ki}} - 1) = U_i(w_{i0}) (x_{ki} - 1) \quad (C.21)$$

Meanwhile differentiation of Eq. (5) gives:

$$\frac{dU_i}{dw_i} = \beta_i e^{\beta_i w_i} = \beta_i U_i = \frac{\ln x_{ki}}{y_{ki}} U_i \quad (C.22)$$

where, Eq. (C.18) is used in the last step.

Substituting from Eqs. (C.21) and (C.22) into Eq. (A.10) then gives the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth as:

$$m_{ki} = \frac{x_{ki} - 1}{\ln x_{ki}} y_{ki} \quad (C.23)$$

This is an alternative but equivalent form to Eq. (37), as can be shown by expanding Eq. (28) as:

$$e^{-\beta_i x_{ki}} - 1 = -(e^{\beta_i y_{ki}} - 1) \quad (C.24)$$

so that, using Eqs. (30) and (C.17)

$$\psi_{ki} - 1 = -(x_{ki} - 1) \quad (C.25)$$

Dividing both sides of this equation by β_i gives

$$\frac{\psi_{ki} - 1}{\beta_i} = -\frac{x_{ki} - 1}{\beta_i} \quad (C.26)$$

Now substituting for β_i from Eq. (33) into the left-hand side and from Eq. (C.18) into the right-hand side of Eq. (C.26) gives the desired result:

$$\frac{\psi_{ki} - 1}{\ln \psi_{ki}} x_{ki} = \frac{x_{ki} - 1}{\ln x_{ki}} y_{ki} \quad (C.27)$$

Eq. (C.23) may be rewritten using Eq. (C.20) as

$$m_{ki} = \frac{F\left(\frac{x_{ki}}{y_{ki}}\right) - 1}{\ln F\left(\frac{x_{ki}}{y_{ki}}\right)} y_{ki} \quad k = W, X \quad (C.28)$$

This may be compared with the expression previously derived as Eq. (38), repeated below and just shown to be equivalent:

$$m_{ki} = \frac{F\left(\frac{y_{ki}}{x_{ki}}\right) - 1}{\ln F\left(\frac{y_{ki}}{x_{ki}}\right)} x_{ki} \quad k = W, X \quad (38)$$

Eqs. (C.28) and (38) will give the same result for every pair of MAP and MAC, (x_{ki}, y_{ki}) .

In the same way that Eq. (19) can be transformed into Eq. (C.15) by replacing y_{ki} by x_{ki} and x_{ki} by y_{ki} , carrying out a similar transformation on Eq. (38) will produce Eq. (C.28).

C.3. Comparing the expressions for m_{ki} for the Positive Exponential and Negative Exponential utility functions

As noted in Section 5, Carthy et al. (1999) suggested that the calculational method for finding the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth used under the Negative Exponential utility function could also be employed to find m_{ki} under the Positive Exponential utility function by transposing x_{ki} and y_{ki} . They are correct in this assessment, since replacing y_{ki} by x_{ki} and x_{ki} by y_{ki} in Eq. (38) will produce Eq. (C.28), and Eq. (C.28) for m_{ki} under the Positive Exponential utility function is identical to Eq. (C.15) for m_{ki} under the Negative Exponential utility function. Moreover, Eq.

(C.15) has been found by replacing y_{ki} by x_{ki} and x_{ki} by y_{ki} in Eq. (19).

However, the comments in Jones-Lee and Loomes (2015) suggest that those authors do not seem to have appreciated the more general point that all four Eqs. (19), (C.15), (38) and (C.28) will produce the same result for every pair of MAP and MAC, (x_{ki}, y_{ki}) . Hence in attempting to find the marginal rate of substitution, m_{ki} , of non-injury probability in place of wealth under the Positive Exponential utility function, it is not necessary to swap the places of y_{ki} and x_{ki} in the Eq., (19), that gives m_{ki} under the Negative Exponential utility function. Eq. (19) can be used directly. As shown in this Appendix, this is a necessary consequence of the model used in the two-injury chained method.

Appendix D. Finding the expected value, $E(M_i)$, of the ratio of Λ_i to its complement

The probability distribution, $q(m_i)$, for m_i may be found as:

$$q(m_i) = f(\lambda_i(m_i)) \left| \frac{d\lambda_i}{dm_i} \right| \quad (D.1)$$

From Eq. (73),

$$\lambda_i = \frac{m_i}{m_i + 1} \quad (D.2)$$

and so:

$$\frac{d\lambda_i}{dm_i} = \frac{1}{(m_i + 1)^2} \quad (D.3)$$

Meanwhile we may use Eq. (73) to define $m_a = m_i(\lambda_a)$:

$$m_a = \frac{\lambda_a}{1 - \lambda_a} \quad (D.4)$$

After using Eq. (D.2) in Eq. (77), we may combine Eqs. (D.1), (D.3) and (D.4) to give the probability distribution, $q(m_i)$, as:

$$q(m_i) = \begin{cases} \frac{2}{(1 + \lambda_a)(m_i + 1)^2} & \text{for } 0 < m_i < m_a \\ \frac{2}{(1 - \lambda_a^2)(m_i + 1)^3} & \text{for } m_a < m_i \leq \infty \end{cases} \quad (D.5)$$

The expected value, $E(M_i)$, may then be found as:

$$\begin{aligned} E(M_i) &= \int_{m_i=0}^{\infty} m_i q(m_i) dm_i \\ &= \frac{2}{1 + \lambda_a} \int_{m_i=0}^{m_a} \frac{m_i}{(m_i + 1)^2} dm_i + \frac{2}{1 - \lambda_a^2} \int_{m_i=m_a}^{\infty} \frac{m_i}{(m_i + 1)^3} dm_i \end{aligned} \quad (D.6)$$

so that:

$$\begin{aligned} E(M_i) &= \frac{2}{1 + \lambda_a} \left(\frac{1}{m_i + 1} + \ln(m_i + 1) \right) \Big|_0^{m_a} \\ &\quad + \frac{2}{1 - \lambda_a^2} \left(-\frac{2m_i + 1}{2(m_i + 1)^2} \right) \Big|_{m_a}^{\infty} \end{aligned} \quad (D.7)$$

Hence:

$$\begin{aligned} E(M_i) &= \frac{2}{1 + \lambda_a} \left(\frac{1}{m_a + 1} + \ln(m_a + 1) - 1 \right) \\ &\quad + \frac{2}{1 - \lambda_a^2} \left(\frac{2m_a + 1}{2(m_a + 1)^2} \right) \end{aligned} \quad (D.8)$$

Setting, $\lambda_a = 0.99$, it follows that $m_a = 99$ and $E(M_i) = 4.63$.

References

- Bernoulli, D., 1738, 1954, "Exposition of a New Theory on the Measurement of Risk", *Econometrica*, Vol. 22, No. 1, Jan., pp. 23–36. Translated by Dr. L. Sommer from "Specimen Theoriae Novae de Mensura Sortis", *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, Tomus V (Papers of the Imperial Academy of Sciences in Petersburg, Vol. V), 1738, pp. 175–192. Available at (<http://www.jstor.org/stable/1909829>) (Accessed February 2014).
- Boadway, R.W., Bruce, N., 1984. *Welfare Economics*. Basil Blackwell, Oxford, United Kingdom.
- Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Spencer, A., 1999. On the contingent valuation of safety and the safety of contingent valuation: Part 2-the CV/SG chained approach. *J. Risk Uncertain.* 17, 187–213.
- Chilton, S., Covey, J., Jones-Lee, M., Loomes, G., Pidgeon, N., Spencer, A., 2015. Response to Thomas and Vaughan. *Process Saf. Environ. Protect.* 93, 293–298.
- Department for Transport, 2013, WebTAG Data Book, WebTAG Table A 4.1.1, 29 October. Available at: (<https://www.gov.uk/government/publications/webtag-tag-data-book>). (Accessed May 2014).
- HSE, 2014, ALARP Suite of Guidance, (<http://www.hse.gov.uk/risk/expert.htm>).
- Hicks, J.R., 1939. *Value and Capital*. Oxford University Press, Oxford UK.
- Jones-Lee, M., Loomes, G., 2015. Final response to Thomas and Vaughan. *Lett. Process Saf. Environ. Protect.* 94, 542–544.
- Kaldor, N., 1939. Welfare propositions and interpersonal comparisons of utility. *Econ. J.* XLIX, 549–552.
- Matheson, J., Summerfield, C. (Eds.), 2000. *Social Trends 30, 2000 Edition*. Office of National Statistics, The Stationery Office, London, UK.
- Nathwani, J.S., Lind, N.C., Pandey, 1997. *Affordable Safety by Choice: The Life Quality Method*. Institute for Risk Research, University of Waterloo, Waterloo, Ontario, Canada.
- Nathwani, J.S., Lind, N.C., Pandey, 2008. *Engineering Decisions for Life Quality. How Safe is Safe Enough?* Springer, Dordrecht, Germany.
- Pandey, M.D., Nathwani, J.S., 2003. A conceptual approach to the estimation of societal willingness-to-pay for nuclear safety programs. *Nucl. Eng. Des.* 224, 65–77.
- Pandey, M.D., Nathwani, J.S., Lind, N.C., 2006. The derivation and calibration of the life-quality index (LQI) from economic principles. *Struct. Saf.* 28, 341–360.
- Popper, Sir K.R. (1934), *Logik der Forschung*, translated as *The Logic of Scientific Discovery*, 1959, revised edition 1980, Hutchinson, London.
- Spackman, M., Evans, A., Jones-Lee, M., Loomes, G., Holder, S. and Webb, H., 2011, "Updating the VPF and VPIs: Phase 1: Final Report [to] Department for Transport", NERA Economic Consulting. Available at: (<http://assets.dft.gov.uk/publications/pgr-economics-rdg-updatingvpfpvi-pdf/vpivpfreport.pdf>). (Accessed May 2014).
- Sunstein, C.R., 2004a. Lives, life-years, and willingness to pay. *Columbia Law Rev.* 104, 205–252.
- Sunstein, C.R., 2004b. Valuing life: a plea for disaggregation. *Duke Law J.* 54, 384–445.
- Thomas, P., 1999. *Simulation of Industrial Processes for Control Engineers*. Butterworth-Heinemann, Oxford, UK (ISBN 0 7506 4161 4. Chapters 2 and 18).

- Thomas, P.J., 2010. An absolute scale for measuring the utility of money. In: Proceedings of the 13th IMEKO TC1-TC7 Joint Symposium Without Measurement no Science, without Science no Measurement. IOP Publishing, City University, London, UK, p. 012039, <http://dx.doi.org/10.1088/1742-6596/238/1/012039> (September 1–3, Journal of Physics: Conference Series 238).
- Thomas, P.J., 2013. The importance of risk-aversion as a measurable psychological parameter governing risk-taking behaviour. In: Proceedings of the 2013 Joint IMEKO TC1-TC7-TC13 Symposium Measurement across physical and behavioural sciences, 4–6 September, Genova, Palazzo Ducale-Italy (Journal of Physics: Conference Series 459). <http://iopscience.iop.org/1742-6596/459/1/012052>.
- Thomas, P.J., Stupples, D.W., Alghaffar, M.A., 2006a. The extent of regulatory consensus on health and safety expenditure. Part 1: development of the J-value technique and evaluation of regulators' recommendations. Trans. IChemE, Part B, Process Saf. Environ. Protect. 84 (B5 (September)), 329–336.
- Thomas, P.J., Stupples, D.W., Alghaffar, M.A., 2006b. The extent of regulatory consensus on health and safety expenditure. Part 2: applying the J-value technique to case studies across industries. Trans. IChemE, Part B, Process Saf. Environ. Protect. 84 (B5 (September)), 337–343.
- Thomas, P.J., Jones, R.D., Kearns, J.O., 2010a. The trade-offs embodied in J-value analysis. Process Saf. Environ. Protect. 88 (3 (May)), 147–167.
- Thomas, P.J., Jones, R.D., Boyle, W.J.O., 2010b. The limits to risk aversion. Part 2: the permission point and examples. Process Saf. Environ. Protect. 88 (6 (November)), 396–406.
- Thomas, P.J., Jones, R.D., 2010c. Extending the J-value framework for safety analysis to include the environmental costs of a large accident. Process Saf. Environ. Protect. 88 (5 (September)), 297–317.
- Thomas, P., Boyle, W., Kearns, J., 2010d. The quantum of wealth. Measurement + Control 43 (5 (June)), 156–158.
- Thomas, P., Vaughan, G., 2013. All in the balance: assessing schemes to protect humans and the environment. Nucl. Future 9 (3 (June)), 41–51.
- Thomas, P.J., Vaughan, G.J., 2015a. Testing the validity of the “value of a prevented fatality” (VPF) used to assess UK safety measures. Process Saf. Environ. Protect. 94, 239–261, <http://www.sciencedirect.com/science/article/pii/S0957582014000962>.
- Thomas, P.J., Vaughan, G.J., 2015b. “Testing the validity of the “value of a prevented fatality” (VPF) used to assess UK safety measures”: Reply to the comments of Chilton, Covey, Jones-Lee, Loomes, Pidgeon and Spencer. Process Saf. Environ. Protect. 93, 299–306, <http://www.sciencedirect.com/science/article/pii/S0957582014001797>.
- Treasury, U.K., 2011. The Green Book, Appraisal and Evaluation in Central Government. The Stationery Office, London (Accessed 2014). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf.
- Wolff, J., and Orr, S., 2009, “Cross-sector weighting and valuing of QALYs and VPFs. A report for the inter-departmental group for the valuation of life and health”, Final Report, 08.07.09. (<http://www.ucl.ac.uk/cpjh/docs/IGVLH.pdf>).